# Desirable Characteristics of an Online Data Commons for Spatially Referenced, Locally Generated Data from Disparate Contributors

*James Campbell and Harlan Onsrud*

***Abstract:*** *A significant body of spatially referenced, locally produced data in small isolated collections exists on the hard drives and backup systems of individual researchers, nonprofit groups, private associations, small companies, universities, and nongovernmental organizations across the United States. From a practical perspective, that data currently is unavailable to professional scientists and to the general public. If there were an online environment where that data could be deposited or registered and readily found, what infrastructure characteristics might potential users find desirable for them to be willing and interested in finding, consulting, and using such data? While there are major national and international initiatives such as the Global Earth Observation System of Systems (GEOSS) that are providing a gateway for access to millions of spatially referenced datasets, primarily from national government data sources, a similar gateway to access spatially referenced, locally produced datasets from disparate private and nonprofit sources has yet to emerge. If one or more were to emerge, what characteristics should be incorporated into the design to make it useful to users of the portal or gateway?*

*Based on data-preservation literature, this study posits three potential characteristics as desirable: make conditions of use of data files clear to potential users; provide a variety of ways to search for data; and enable users to access comments and feedback from prior users, and add comments of their own. These three characteristics were examined because they often are not provided or inadequately provided in general-purpose portals for finding geographic data and services. A combination of qualitative and quantitative methods was used and the results of the analysis using both methods support the hypothesis.*

## INTRODUCTION

### Background

A significant body of spatially referenced, locally produced, small-scale data developed for specific local purposes exists on the hard drives and backup systems of individuals, nonprofit groups, private associations, universities, private companies, and other nongovernmental organizations across the United States. Spatially referenced data, as the term is used here, is data that refers to a particular physical location. Examples might include a university botany class project that locates and catalogs all the trees more than 15 feet tall in a small town; a homeowners' association that monitors the water quality and plant growth of the lake on which members' properties are located; a land trust that records environmental easements; or a historical museum that ties its photographic images to their physical locations, among many others.

In all these cases, the data gathered by these small local originators could be of great value to others if its existence were known. At present, however, very little of this data is available from a practical perspective to other scientific researchers and potential users. It is, for all intents and purposes, completely or partially "invisible."

While much emphasis has shifted in recent years to providing geospatial services, there still is a strong need for service developers to be able to find and exploit existing geographic data that would make those services more effective and efficient. Many efforts at the national and state levels are being made to make government-generated spatially referenced data available to the public. In the United States and in other countries around the world, initiatives are under way to make geographic information more freely available to scientists and to the general public. In English-speaking countries, for example, UK Location (http://location.defra.gov.uk) in the United Kingdom, the Atlas of Canada (http://atlas.gc.ca/site/english/index.html), and Geoscience Australia (www.ga.gov.au) provide open access to some government-generated spatially referenced data. In the United States, initiatives such as the National Map (http://nationalmap.gov), the National Atlas (www.nationalatlas.gov), and the geospatial section of data.gov (http://www.data.gov/geospatial/) serve similar functions. These U.S. sites contain a wider array of data than many other national portals because the U.S. federal government cannot hold copyright on materials it generates, and because some state governments make their state-level data visible through these gateways. Efforts also are under way to make international sharing of large datasets more viable, especially with regard to divergent approaches to data licensing and use rights (Onsrud et al. 2010). GEOSS Data

Collection of Open Resources for Everyone (GEOSS Data-CORE 2014) is an example of an international initiative to support open access to geographic data gathered by governments across nine societal benefit areas (GEOSS 2014).

Similarly, disciplinary and special purpose repositories exist to capture large sets of spatially referenced data. Examples include PANGAEA (http://www.pangaea.de) and OneGeology (http://www.onegeology.org).

Google Maps, Google Earth, Virtual Earth, and Open Street Maps provide structured environments where the user may take advantage of a data-gathering and display infrastructure to contribute data or volunteer effort to a commercial or open-data environment. In these information infrastructure environments, legal and data management issues as well as data format issues are closely controlled by the infrastructure system provider. These are not infrastructure environments for depositing or finding diverse geographic datasets, and this article does not address such environments.

We conclude that no gateway exists analogous to the Global Earth Observation System of Systems (GEOSS) that could provide more visible and efficient access to millions of spatially referenced datasets drawn from disparate locally generated sources. Note that the GEOSS is a *portal* or *gateway* for finding relevant geographic data and services rather than a *repository* of geographic data itself. Furthermore, the metadata on geographic data and services contained within the GEOSS is provided or mined from primarily national and international government members and participating organizations of the Group on Earth Observations (GEO). The GEOSS serves as an exemplar of the kind of infrastructure that can make geospatial data files and services from widely disparate cooperating sources much more readily findable.

## VOLUNTEERED GEOGRAPHIC INFORMATION (VGI)

In the past decade, regular people have become producers as well as consumers of geospatial data, a phenomenon variously called neogeography (Turner 2006, Sui 2008), ubiquitous cartography (Gartner et al. 2007), collaboratively contributed geographic information (Bishr and Mantelas 2008), and volunteered geographic information (Goodchild 2007). VGI seems to be the most widely used term at present.

Affordable, portable GPS devices have made it possible for anyone to make a quite accurate observation of the position of an object on the face of the earth. Simple-to-use infrastructures that use Google Maps, Open Street Maps, or similar frameworks make it easy to add those observations to a map, and to attach notes or information to the location. To date, the great bulk of VGI activity has involved this form of adding locations and labels of features within a mapping facilitation framework or to already existing maps. At the observation level, then, VGI contributors can contribute data in many situations as well as trained geographers could in pre-GPS days.

Adding or correcting locations, names, and characteristics of features on a map base such as Google Maps or Open Street Maps is a type of spatially referenced data but there are many other types including complete datasets of various kinds such as the examples mentioned previously. Most of the examples involve "asserted" rather than "authoritative" data (Bishr and Mantelas 2008). In VGI-contributed environments, where disparate datasets are only asserted as potentially useful and not vouched for, context becomes crucial. VGI data, or any data, collected for one specific purpose may not be relevant or useful or even accurate for a different purpose. Potential online environments that may feature collections of data generated locally for disparate purposes need to contextualize that data for the data to be useful.

## DESIRABLE CHARACTERISTICS OF AN ONLINE SPATIALLY

### Referenced Data Repository

Simply having an online gateway or home for widely disparate, spatially referenced, locally generated datasets could be of significant use for providing access to this type of data. It probably would be of greatest use to geospatial specialists and professionals desiring to find and draw from existing spatially referenced data to provide further products and services. We refer to this perceived online gateway or home as a Commons of Geographic Data (CGD). However, if such a facility or capability, centrally located or distributed, is to be of maximal use over time to both professional scientists and to interested nonprofessionals, a number of studies and reports suggest that it should include functionality that enables users to know usage rights and search for and discover data using standards-based metadata, and provide users with a way to access evaluation commentary from previous users of the datasets and offer comments of their own. See these common elements in, for example, Report of the Workshop on Opportunities for Research on the Creation, Management, Preservation and Use of Digital Content (IMLS 2003), Licensing Geographic Data and Services (NRC 2004), and To Stand the Test of Time: Long Term Stewardship of Digital Data Sets in Science and Engineering (ACRL 2006).

In a commons-type environment for data users, data is made available under a license—if a license is necessary to use the data—that grants permission for use as long as any stipulated conditions are adhered to. This makes it possible for potential users to be sure that they may use any data found in such a commons environment without seeking additional permission from the owner. In such environments, permission already has been granted as long as any conditions specified in the license are respected. Creative Commons licenses are one example of so-called "some rights reserved" license types typically found in a commons environment for materials that are not in the public domain. Creative Commons licenses currently are used in more than half a billion digital works. Creative Commons and its affiliate, Science Commons, have designed several licenses specifically

applicable to datasets (Creative Commons 2014) that could be used in a Commons of Geographic Data.

An online Commons of Geographic Data with the characteristics listed previously does not exist at present. If such an environment were contemplated as a future project, based on the reports previously cited, important questions arise almost immediately. If there were such an online data commons repository for small, privately generated datasets, would people who are interested in spatially referenced data be willing to access and use the data in such a repository? What type of functional characteristics of such a repository or gateway would help to motivate those potential data users to actually examine and possibly use the data located there for their own purposes?

It may seem reasonable to assume that such characteristics would be desirable to potential users, but at this point in time, reasonable or not, this still is an assumption. The goal of this research is to address this question empirically.

## HYPOTHESIS

The purpose of this research is quite practical. It is hoped that the results may provide some guidance for future architects of an online Commons of Geographic Data about functionality that potential users would be interested in finding in an online commons environment for spatially referenced small datasets from disparate sources, if and when such a commons environment is constructed. The results could suggest several areas for future research, and might also be of use to those who currently operate data gateways or repositories that they would like to make more responsive to users' interests.

Based on common elements in the reports noted previously as well as in other data-preservation related studies (e.g., Committee on Science, Engineering, and Public Policy (U.S.) 2009, Interagency Working Group on Digital Data 2009), we hypothesized that potential data users would be willing to consider using data accessed through an online gateway or data repository if such a facility included:

(a) a simple, clear licensing mechanism that reveals ownership of, and conditions for use of, the contributed data;
(b) a simple, effective searching/finding mechanism that provides an option to search using either *Thesaurus*-controlled vocabulary, "plain English" keywords, or location; and
(c) a simple postpublication peer-evaluation mechanism that will provide information on quality and suitability for purpose for users.

## METHOD

To test this hypothesis, we used a combination of qualitative and quantitative research procedures (Onwuegbuzie and Leech 2004; Ragin, Nagel, and White 2004). Personal interviews were conducted with ten people who were regular users of spatially referenced data. These particular interviewees also were generators of spatially referenced data. The findings from these qualitative interviews were used to construct an online questionnaire, and results from that questionnaire with responses from a much larger group (139 people) were compared with the results from the interviews to see if the qualitative results were supported by quantitative data.

## METHODOLOGICAL LIMITATIONS

The respondents in this study are not in any way meant to be considered a statistical or otherwise representative sample of potential data users of an online commons gateway or repository for spatially referenced datasets from disparate sources. The major reason for not attempting to select a representative sample of potential users is that the universe of such users is unknown and probably unknowable. Thus, the combination of qualitative in-depth interviews with quantitative data was chosen to produce findings that would be informative, even though not "proven" in a statistical sense, for future designers of an online commons-type geospatial data environment, and that could suggest directions for future study.

All participants in the study were self-selected. In addition, to generate quantitative responses online, given the reverse traceability of personal user information in today's online environment, potential respondents were guaranteed anonymity by requesting no geographic, employment, or other demographic information. This makes some types of statistical analysis impossible.

## INTERVIEWEES AND DATA TYPES

Interviewees were selected based on a "snowball technique" (Maxwell 2005). Interviewees were referred by word of mouth from those interested in spatially referenced data who were located in geographic areas accessible to the authors. Those who agreed to participate were asked if they could recommend others who might be potential interviewees. In the final group of ten interviewees, seven were from Maine, one from Massachusetts, one from Pennsylvania, and one from North Carolina.

One interviewee was a graduate student working on a spatial-data research project; one regularly dealt with spatially referenced data as part of the respondent's employment, although the role the respondent held in this study was as a volunteer citizen on a municipal committee. About half the respondents were familiar with and used GIS software to a greater or lesser degree; about half did not. Four were involved with land trusts of one type or another, one was an author of nature books, one a high school teacher, one a local museum curator, and the others were involved with other types of local civic groups. All the spatially referenced data that these originators were gathering were deemed by the investigators and the gatherers to be of potential interest to others in the future but none of the data was available on the Web.

## QUALITATIVE DATA-COLLECTION PROCESS

The purpose of these qualitative interviews was to test whether the hypothesis above would hold, and to discover if other important

desirable characteristics arose spontaneously in the interviews. All interviews were conducted from the same interview instrument by the same interviewer. The interviews were transcribed and coded, and then the transcripts were checked against the voice recordings for accuracy. A summary of key points then was sent to each interviewee for correction, if necessary, and for confirmation. None of the interviewees who responded submitted any corrections other than spelling errors.

Because all interviewees were asked the same set of questions, initial top-level codes were based on those questions, e.g., "conditions" (which owners might put on use of contributed data); "metadata" (short description, keywords, search order, etc.); "evaluation" (valuable or not, amount of time willing to spend commenting, etc.). As additional aspects of responses appeared, subcategories for the major categories were added to make meanings more precise, and a few additional top-level codes added for topics that emerged that were not specific responses to asked questions but that were relevant to overall online data commons use.

## QUANTITATIVE DATA-COLLECTION PROCESS

Based on the information generated in the analysis of the qualitative data, an online questionnaire was constructed to see if others who identified themselves as users of spatially referenced data would agree with the responses of the ten interviewees regarding the hypothesis points. Notice of the existence of the questionnaire along with an invitation to participate in the research was sent out to listservs of those concerned with geographic information of different types, specifically to members of the Global Spatial Data Infrastructure Association and to members of the Maine Geolibrary listserv. In addition, printed flyers inviting participation were distributed at a conference of the Maine GIS User Group and the Maine Municipal Association.

The survey instrument used the first question to separate those who were owners of, or who had significant influence on data sharing in their organizations (potential contributors), from those who considered themselves only potential data users.

All those who identified themselves as potential contributors also considered themselves potential users, and there were additional respondents who considered themselves users only. We report on the results of the questions answered by all users, including those who also identified themselves as owners or controllers of spatially referenced data. There were 11 questions data users were asked to answer in the survey, of which three requested text-based responses.

As in the qualitative portion of the research, no attempt was made to construct a statistically valid sample. Rather, the goal was to gather a reasonable number of responses from self-identified potential users of spatially referenced data to either support or invalidate the qualitative research findings.

There was a total of 197 click-throughs from the survey splash page to the actual survey instrument. Each click-through response was given a specific ID for analysis purposes. Of 197 click-throughs, 139 completed some or all of the questions put to users.

## RESULTS

We review the results by each hypothesis subpart. Although the prior discussion refers to both portals and repositories for geographic data, with the human subjects we focused on the simpler concept of data repositories. However, we believe the results are generalizable for also guiding feature developments for portals or gateways such as GEOSS that lead to distributed repositories or portals.

## HYPOTHESIS SUBPART (A): SIMPLE CLEAR TERMS OF USE

Hypothesis: Data users would be willing to consider using data in an online data repository if such a repository included a simple, clear licensing mechanism that reveals ownership of, and conditions for use of, the contributed data.

## QUALITATIVE RESULTS

All ten of the interviewees indicated that they would want to be able to check license conditions before they decided to download and use data, and that they would respect any conditions that were put on the use of the data in a particular file. Most indicated that they would want a simple-to-understand statement of what they could or could not do with a data file. In the words of one interviewee: "I would want to be able to identify the conditions or at least get a sense of the conditions very quickly . . . I am not going to spend a lot of time reading a three-page license agreement."

Several assumed that any conditions for use would be stipulated when a file was found, and certainly by the time it was opened, although another interviewee said that the interviewee always scans the Web page a file appears on to see if, for example, attribution is required.

Several interviewees referred to ethical considerations when describing whether and why they would check any licensing conditions before using the data in any but a personal way. Two of the interviewees indicated specifically that they would not bother to check for licensing conditions if they were just looking at the data for their own information, but if they contemplated using it in any additional way, they would check and respect any conditions of use.

Interviewees were asked if the presence of conditions of use that were clearly stated before opening a file might impact whether they would choose to look at a data file or not. Responses were evenly divided between those who would look at the data anyway and those who would not bother if they felt the conditions would preclude the use that they might wish to put the data to.

## QUANTITATIVE RESULTS

Results from responses to the online questionnaire are consistent on this topic with those gleaned from the personal interviews.

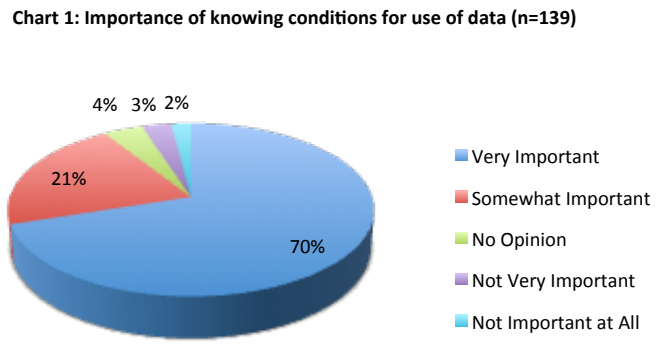



**Chart 1.** Importance of knowing conditions for use for data (n=139)

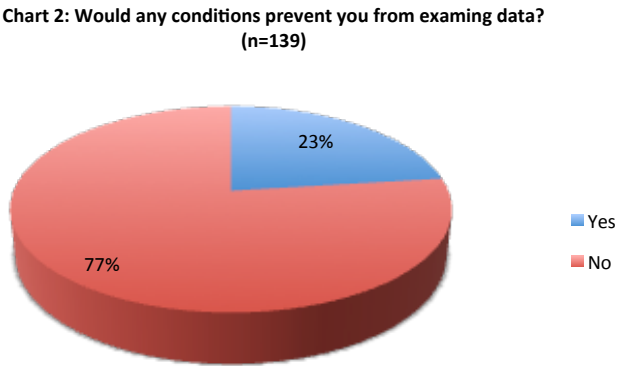



**Chart 2.** Would any conditions prevent you from examining data? (n=139)

Users were asked in each question "If you were looking for data that others had contributed to an online commons-type environment, please indicate how important each of the following would be in your decision of whether to access and/or use such data . . ."

Users were given five choices:

- Very Important
- Somewhat Important
- No Opinion
- Not Very Important
- Not Important at All

This first question asked how important it would be that "Conditions for the use of the data are clear." See Chart 1. (Note that all the following chart percentages are rounded.)

The importance of knowing the conditions for use expressed by interviewees is mirrored in the larger population of questionnaire respondents, with 91 percent indicating that such knowledge would be "Very Important" or "Somewhat Important" to them

Addressing the question of whether licensing conditions put on the use of the data would affect potential users from accessing the data, respondents were asked: "If conditions for use of the data were clear, e.g., requiring attribution or noncommercial use only, might there be any conditions that would prevent you from examining the data?" (See Chart 2.)

Of those questionnaire respondents who responded "Yes" to this question, examples of conditions that might prevent users from examining a data file varied. The predominant response concerned limitations on commercial use. Some other reasons included cost, administrative requirements, concern about data quality, limited bandwidth that would preclude downloading large files, and inability to modify the data for their own use.

## HYPOTHESIS SUBPART (B): SEARCH MECHANISM

Hypothesis: Data users would be willing to consider using data in an online data repository if such a repository included a simple, effective searching/finding mechanism that provides an option to search using either Thesaurus-controlled vocabulary, "plain English" keywords, or location.

## QUALITATIVE RESULTS

None of the interviewees said that they would search for data based on Thesaurus-controlled vocabularies. All would begin searches using either natural language keywords and phrases, or location terms. All interviewees indicated that they might use either strategy first depending on what they were looking for at a particular time. About half indicated that they usually would begin with topic keywords, about half with location. However, each group then would use the other strategy to help narrow their results.

For example, an interviewee who served on a municipal recreation committee interested in resident uses of lakes described a strategy for finding that type of information: "So when we start to look out and search the Internet we throw a broad net at the beginning based on certain things like those lake management plans but when we get down to specifics we start looking at information of lakes that are more in the same latitude or in close proximity to where the municipality that we live is." Another interviewee who worked with a local land trust took a different approach: "In terms of my work and the way I would do it, it would be place based; it would be coming from the place to the information."

In either case, interviewees found being able to begin their searches either by topic or place keywords was important for their search strategies.

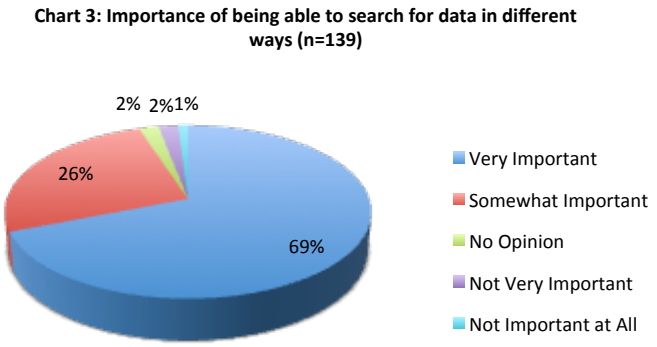



**Chart 3.** Importance of being able to search for data in different ways (n=139)

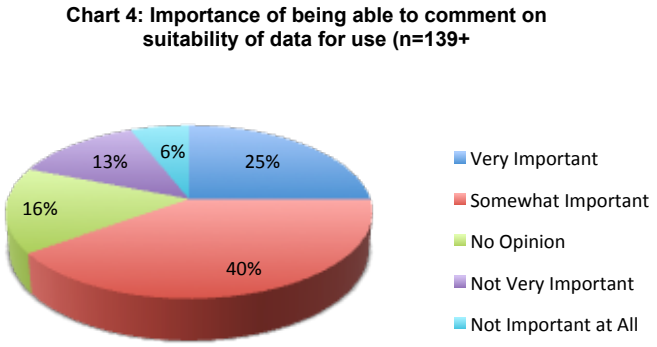



**Chart 4.** Importance of being able to comment on suitability of data for use (n=139)

## QUANTITATIVE RESULTS

Questionnaire respondents were asked how important the "Ability to search for data in different ways, e.g., by location, keyword, etc." would be to them. The results are consistent with those from the interview phase of this research. (See Chart 3.)

Being able to conduct searches using different starting points, including location and natural language keywords, appears to be an important functional capability for an online repository for locally generated, small-scale spatially referenced data.

## HYPOTHESIS SUBPART (C): PEER EVALUATION

Hypothesis: Data users would be willing to consider using data in an online data commons environment if such an environment included a simple post-publication peer-evaluation mechanism that would both provide feedback for contributors, and provide information on quality and suitability for use for users.

## QUALITATIVE RESULTS

In this age of Amazon and online shopping, it is no surprise that interviewees used online shopping comments as an analog to looking at comments/evaluations in an online commons environment for spatially referenced data. Half of the interviewees made comments similar to this one: "I mean I buy CDs on Amazon.com" that indicated familiarity with commercial online retailer commenting systems that they found useful, and indicating that they would consult peer comments and evaluation of data files if such comments were available.

Half of the respondents, however, said that they would look at the data themselves if it were data that might suit their needs, no matter what the comments said. Two indicated that they would look at the data first and only subsequently consult other user comments to see if those corresponded with their own judgments.

Only one interviewee said that the interviewee would be unlikely to consult comments made by others because the interviewee preferred to form a personal opinion directly from the data.

One interviewee indicated that "junk comments" were always a potential problem in evaluation systems and recommended that any such system have a moderator who would screen comments for civility, relevance, and, if possible, quality before posting them.

Other interviewees who would consult comments made by others indicated that while they would not view it as necessary, they would prefer to know who the commenter was so that they could form an opinion about the relevance or quality of the comment source if the commenter were known to them.

Nine of the interviewees indicated that they would be willing to make comments if they felt that they had something useful to say about a file. Most said that they would be willing to spend a limited amount of time, 5 to 15 minutes, to input a comment if there were a simple way to do so.

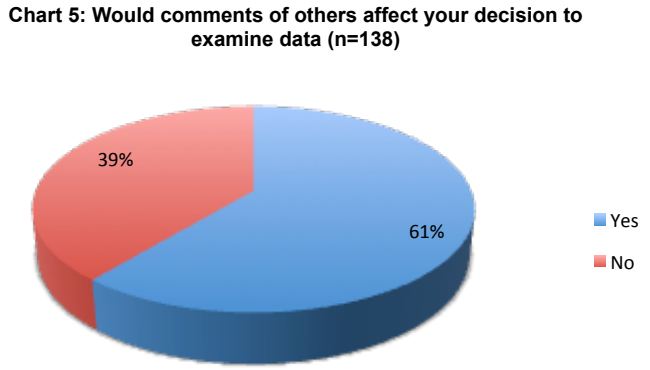
Consistent with the desire to know who made a comment, all nine said that they would be willing to use their own names rather than to use a screen name in offering a comment.

In summary, the majority of interviewees would find a commenting/evaluation system valuable in an online commons repository.

## QUANTITATIVE RESULTS

Support for the "Ability to comment on the suitability of the data for your uses" was not so strong among survey respondents as among interviewees, although it was substantial, with 65 percent finding that capability "Very Important" or "Somewhat Important." (See Chart 4.)

The amount of time that survey respondents would be willing to spend providing a comment generally mirrored what most interviewees would spend, 5 to 15 minutes. Given 139 responses rather than 10 as in the personal interviews, however, it is not surprising that there were a few outliers who would commit any-



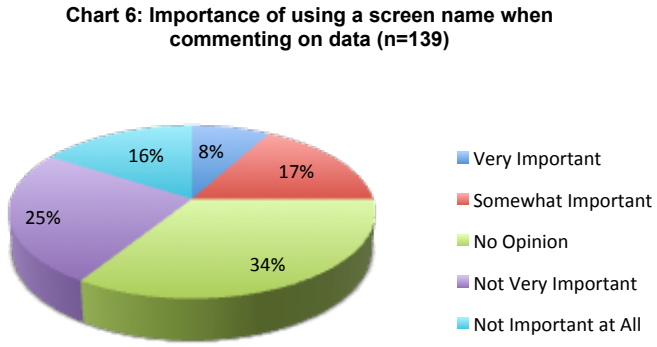




**Chart 5.** Would comments of others affect your decision to examine data? (n=138)

**Chart 6.** Importance of using a screen name when commenting on data (n=139)

where from "no time" to "as much as would be needed."

In response to the question "Would the comments of other users affect your decision about whether to examine data that is available in the repository?" of 138 responses, 61 percent replied "Yes" and 39 percent said "No" (see Chart 5).

When asked to "explain how comments of others might affect your decision about whether to examine data further," a large majority of those who answered (78 of 84) cited comments that dealt with data quality and accuracy. Here, again, the analogy of online commerce sites came up: "Same as eBay. If someone says the data are junk, I'll probably be reluctant to use them."

The other major reason expressed by respondents was not the quality of the data itself but rather the lack of suitability for purpose, e.g., "how the data fits with my base maps."

The "Ability to use a screen name rather than your actual name when commenting" was more of an issue to survey respondents than it was with the interviewees. (See Chart 6.)

While nine of ten interviewees would use their own names rather than a screen name when making comments and preferred to know the identity of those making comments when possible, 25 percent of questionnaire respondents felt it would be "Very Important" (8 percent) or "Somewhat Important" (17 percent) to be able use screen names when commenting, and a third did not express any opinion. The reason for this divergence from the attitudes of interviewees is not explainable based on the data this research gathered. The location of the questionnaire respondents might be an issue for commenting using one's real name, or employment status, or some other variable for which this research did not gather any data.

## SUMMARY AND CONCLUSIONS

This research, subject to the caveats listed below, empirically suggests that it would be desirable from the perspective of potential users of spatially referenced data in an online commons-type environment to provide infrastructure capability that would:

- make conditions of use of files clear to potential users,
- provide a variety of ways to search for data, and
- enable users to access comments and feedback from prior users, and to add comments of their own.

There are other desirable features of a commons-type online infrastructure, as the reports cited previously outline. This research addressed only these three.

## LIMITATIONS

As noted earlier, this research has several limitations that prevent any assertion that the hypothesis is "proven" in the usual meaning of that term. However, we can assert that the hypothesis is supported by the results of this study.

These limitations do not, we feel, limit the usefulness of the research results for their intended purpose: to provide guidance to those who may in the future choose to construct an online commons environment for locally generated, small-scale spatially referenced data that anyone, nonprofessional and professional alike, can use.

## DIRECTIONS FOR FUTURE RESEARCH

This research is based on interviews and on online questionnaire results. Results from the interviews generally are confirmed by the survey results. Although percentages differed slightly, opinions about the hypotheses generally were shared both in the interviews and in the survey responses.

However, there was a noticeable disparity in the perception of the importance of being able to use a screen name rather than a real name to make comments, although because a large number of questionnaire respondents expressed "No Opinion," it is difficult to tell if the disparity was important. The absence of demographic, employment, or geographic location information for interviewees

and questionnaire respondents makes it impossible to explain that divergence based on those characteristics. This is an area in which additional research may be fruitful.

This study made no effort to directly ask comparative questions, e.g., is one factor, such as clarity of conditions, more important than another to respondents? Answers to such questions may be inferred from the responses in the importance respondents placed on each factor, but it also could be desirable to ask comparative questions directly.

## POSSIBLE WIDER APPLICATIONS

While this research focused on a possible future online commons-type environment for spatially referenced data from widely disparate sources, the results could be of some use to operators of existing online spatial-data services. Understanding what is desirable to users in approaching data with which they are not familiar, especially non-GIS professionals, could be helpful for existing services to, for example, make clear in an obvious way any restrictions on use of their data. Portals that do not presently enable users to search for data in different ways may wish to evaluate whether such functionality would be desirable to their existing user base, and whether it might help to increase usage among current nonusers of their services. Sites that do not offer commenting capability may wish to investigate if that functionality might increase usage.

For designers of potential future online environments for spatially referenced data, which might include, for example, university libraries or state library systems, and possibly for operators of existing portals as well, we hope this research, though not designed to be statistically "proven," offers some empirical insight into what online characteristics users find valuable for spatially referenced data repositories and/or portals.

## About the Authors

**James Campbell** (james.j.campbvell@maine.edu) is a PhD Candidate, and Harlan Onsrud (harlan. onsrud@maine.edu) is a Professor in Spatial Informatics, School of Computing and Information Science, University of Maine, 5711 Boardman Hall, Orono, ME 04469. Campbell's research interests focus on access to information from both a technical and a law and policy perspective. Onsrud's research focuses on the analysis of legal, ethical, and institutional issues affecting the creation and use of digital databases and the assessment of the social impacts of spatial technologies.

## References

ARL Workshop on New Collaborative Relationships: The Role of Academic Libraries in the Digital Data Universe. Friedlander, A., and P. Adler, Association of Research Libraries & National Science Foundation (U.S.). 2006. To stand the test of time: Long-term stewardship of digital datasets in science and engineering. Washington, D.C.: Association of Research Libraries.

Bishr, M., and L. Mantelas. 2008. A trust and reputation model for filtering and classification of knowledge about urban growth. GeoJournal 72.

Committee on Science, Engineering, and Public Policy (U.S.). 2009. Ensuring the integrity, accessibility, and stewardship of research data in the digital age. Washington, D.C.: National Academies Press.

Creative Commons. 2014. Http://wiki.creativecommons.org/Data.

Gartner, G., D. Bennett, and T. Morita. 2007. Toward ubiquitous cartography. Cartography and Geographic Information Science 34: 247-57.

GEOSS Data Sharing Action Plan, as accepted at GEO-VII, November of 2010, https://www.earthobservations.org/documents/geo_vii/07_GEOSS%20Data%20Sharing%20Action%20Plan%20Rev2.pdf.

Goodchild, Michael F. 2007. Citizens as sensors: The world of volunteered geography. GeoJournal 69: 211-21.

Interagency Working Group on Digital Data. 2009. Harnessing the power of digital data for science and society. Washington, D.C.: National Science and Technology Council, Executive Office of the President.

Institute of Museum and Library Services (U.S.). 2003. Report of the workshop on opportunities for research on the creation, management, preservation, and use of digital content. Washington, D.C.: Institute of Museum and Library Services.

Maxwell, J. A. 2005. Qualitative research design: An interpretive approach, 2nd Ed., Vol. 41, Applied Social Research Series. Thousand Oaks, CA: Sage Publications.

National Research Council (U.S.) Committee on Licensing Geographic Data and Services. 2004. Licensing geographic data and services, Washington, D.C.: National Academies Press.

Nielsen, J. 2000. Designing Web usability. Indianapolis: New Riders.

Nielsen, J. 1993. Usability engineering. Boston: Academic Press.

Onwuegbuzie, A. J., and N. L. Leech. 2004. Enhancing the interpretation of "significant" findings: The role of mixed methods research. The Qualitative Report 9(4): 770-92.

Onsrud, H., et al. 2004. Public commons of geographic data: Research and development challenges. In Egenhofer, M. J., C. Freska, and H. J. Miller, Eds. Geographic information science, Berlin: Springer-Verlag.

Onsrud, H., and J. Campbell. 2007. Big opportunities in access to "small science" data. Data Science Journal 6, http://www.jstage.jst.go.jp/browse/dsj/6/0/_contents/2.

Onsrud, H., J. Campbell, and B. van Loenen. 2010. Towards voluntary interoperable open access licenses for the Global Earth Observation System of Systems (GEOSS). International Journal of Spatial Data Infrastructure Research 5: 194-215.

Ragin, C., J. Nagel, and P. White. 2004. Workshop on scientific foundations of qualitative research, Washington: National Science Foundation.

Sui, D. 2008. The Wikification of GIS and its consequences; or Angelina Jolie's new tattoo and the future of GIS. Computers, Environment and Urban Systems 32: 1-5.

Turner, A. 2006. An introduction to neogeography. Sebastapol, CA: O'Reilly Media.

Veen, J. 2001. The art and science of Web design. Indianapolis: New Riders.