The Digital Divide and Global Spatial Data and Users

Karen D. KLINE, Japan

Key words: data distribution policy, digital divide, Global Map, data access

SUMMARY

The digital divide is defined as the gap between the haves and have-nots of the technological age, and includes not just the technology, but the requisite infrastructure, such as electrical power and communication lines. It also includes the education required to utilize the Internet effectively. It is argued here that a digital divide exists within the global geospatial data community as well; with those with easy access and the education to productively utilize global environmental data sets on one end of the digital divide spectrum and those with little or no access or education, making use nearly impossible, on the other end of the spectrum.

A review of the users of several data sets (Global Map Versions 0 and 1 from GSI Japan, Global Map Australia from GeoSciences Australia, the Global Land Cover Characteristics Data from EROS Data Center USA, and UNEP/GRID Arendal and Geneva) was conducted. In addition to summary statistics, the self-organizing map algorithm and ordinary least squares techniques were used to analyze the user data.

Based upon these techniques, one point becomes clear: All data are local. Users are more interested in locating and utilizing data that reference geographical locations near them. In addition, several other issues are raised:

- Education is a necessary prerequisite for the appropriate use of global environmental data sets.
- Identification of current non-user communities that may benefit from use of these data should be done.
- Electrical and communications infrastructure are necessary for accessing and utilizing the data, and are not necessarily well established every where yet.
- Data provision should be made in other languages besides English.

The Digital Divide and Global Spatial Data and Users

Karen D. KLINE, Japan

1. INTRODUCTION

The digital divide is simply defined as the gap between the haves and have-nots of the technological age (Balakrishnan 2001; Dasgupta, Lall et al. 2001; James 2003). However, this definition includes not just the technology, but also the requisite infrastructure, such as electrical power and communication lines, and education to utilize the Internet effectively. The digital divide, as defined here, exists in many different places. For example, not only between countries, but within countries as well, such as the urban/rural divide often seen in the U.S. when it comes to Internet and telephone connections (Department of Commerce 1999). But it also exists within and between groups. Not all universities around the world are equal when it comes to technological capabilities and Internet access.

It is argued here that a digital divide exists within the global geospatial data community as well; with those with easy access and the education to productively utilize global environmental data sets on one end of the digital divide spectrum and those with difficult or no access or education, at the other end of the spectrum.

To demonstrate the existence of this digital divide within the global data user community, the users of several data sets were studied, using the self-organizing map algorithm and ordinary least squares. The data set users examined include Global Map Versions 0 and 1 served from the Geographical Survey Institute (GSI), Tsukuba, Japan, Global Map Version 1 for Australia served from GeoSciences Australia, Canberra, the Global Land Cover Characteristics (GLCC) data served from the U.S. Geological Survey's EROS Data Center in Sioux Falls, South Dakota, and the United Nations Environment Programme (UNEP)/Global Resource Information Database (GRID) centers in Arendal, Norway and Geneva, Switzerland.

2. METHODOLOGY

This section provides a brief overview of the data collection and collation methods, as well as a summary of the self-organizing map algorithm.

2.1 Data Collection

Several assumptions were made related to differences in the data sites in the interest of establishing a methodology for investigating the users of global data sets. First, the six data providers are not equal. For example, the UNEP/GRID sites do not provide a single data set, but rather a collection of data. Second, the information collected was not consistent, and information occasionally needed to be inferred from the information submitted. Third, English is not the first language of many of the users, and their responses to questions occasionally were difficult to decipher. In addition to which, my understanding of what they wrote may not be the same as what they meant to write. Fourth, there was no quality control in data collection efforts. User information is collected on an ad hoc basis, and is based on the willingness of data users to provide accurate information. The assumption was made that members of the scientific, academic, and educational communities and others interested enough in downloading the respective data sets would provide accurate information.

The assumption was also made that the user information collected from these data providers is the population of users for these particular data sets. This does not take into account the possibility that someone downloaded the data, used it, then passed it along to someone else (the "National Geographic" effect – publish 20,000, but 200,000 people actually see the magazine (personal communication, Goodchild, 2004)). While the general data using population is larger (many data sets are available), the population here is defined as those users who downloaded either UNEP/GRID Arendal, UNEP/GRID Geneva, GLCC, Global Map Versions 0 or 1 (from Tsukuba and Canberra) and have an entry in one of the user data collection systems.

What was not addressed was whether or not the global data and the resolution or scale (1:1km or 1:1,000,000) were appropriate for the application the user indicated they were going to use the data for. Without actually "following" the user and finding out whether or not they used the data as they stated they would, it is impossible to answer this question. To attempt to find out whether or not users are using the data at all, follow on surveys sent to each individual user would be needed.

Table 1: Fields in the User Database

Variable	Description
USER	the unique identifying code assigned to each individual
DATA SET	the name of the data set they downloaded (Arendal, GLCC, or Global Map, Australia, GM Version 0)
DATE	the date they registered in yyyymmdd format
COUNTRY	the country they resided in when they downloaded the data set
CONTINENT	the continent or region where the country is located
DOMAIN	the user's domain (i.e., government, education, etc.)
APPLICATION	the application for which the dataset will be used
LOC COUNTRY	the country location for their research focus
LOC CONT	the continent or region in which the research site is located
SCALE	whether their interest is at the local (country level), regional, or global scale

Table 1 provides a summary of each of the fields in the user database, from which all subsequent work was done. The continent break-down was determined based upon the information in the CIA World Factbook (Central Intelligence Agency 2002). The domain closely mirrors the actual location of the user in terms of their work location, i.e., academic (graduate level education and research), education (K-12 and community college education), government, non-governmental organization, etc. 'Personal' was used to indicate those users not affiliated with a clearly identifiable domain or that have an account with AOL.com or Yahoo.com or their equivalent. The assigned application code was based upon those first developed for the Potential Commercial Applications of EOS data project (Hadley, Estes et al. 2000); a summary list is provided in Table 2. Several additional categories were added, included Research (RES) for those entries that were clearly research, but more information was not available; Reference (REF) for those users that indicated they were interested in the map to look at or to find a location.

Table 2: Application Codes

Application Category	**Code**
Agriculture	AGR
Air quality	AIRQU
Economic development and conservation	EDC
Emergency management	EMERG
Fisheries	FISH
Forestry	FOREST
Geology	GEOL
Information and intelligence	INFO
Land use and land cover	LULC
Mapping, charting, and geodesy	MAP
Marine	MARINE
Media, press, and education	MPRED
Public health	PUBHLTH
Rangeland	RNGLND
Recreation and tourism	REC
Transportation	TRANSP
Urban and regional planning	URBPL
Water quality	WATQU
Water resources	WATRE
Weather and climate	CLIM

From the individual user database compiled from the raw user data, a second data set was derived, which contains country level information. Additional data variables were added, including socioeconomic data (Central Intelligence Agency 2002), Internet access and host data (Central Intelligence Agency 2002)Internet Systems Consortium (www.isc.org)), and aid flow data (Organisation for Economic Co-Operation and Development 2002). The final data set included, for each country listed in the CIA World Factbook, the number of users of each data set, a breakdown by application category as well as domain category, and the other variables from previously mentioned sources.

2.2 Self-Organizing Map

The Self-Organizing Map handles multivariate data very well, and requires no a priori knowledge of the data set being input into the algorithm. To this end, it was chosen as one method to explore and analyze the user data.

The Self-Organizing Map (SOM) algorithm is defined as an unsupervised learning neural network algorithm that takes a set of input data, most often multivariate, and "organizes" the data into a 2-dimensional grid of nodes, or "tessellation" (Skupin 2003), with similar items closer together than dissimilar items (Kohonen, Kaski et al. 2000; Kohonen 2001). The Self-

Organizing Map algorithm lends itself quite well to visualization of multivariate, or high-dimensional data (Kangas and Kohonen 1996; Vesanto 1999; Skupin 2003). In general, any set of data for which a similarity or dissimilarity measure can be calculated for each of its members to each other can be run through a SOM (Kohonen and Somervuo 2002). In addition, “no a priori assumptions about the distribution of the data need to be made” (Deboeck 1998). Kohonen also notes that the SOM is an appropriate tool for visualizing “natural” data, “their distributions are non-Gaussian; and their statistics are nonstationary” (Kohonen 2001).

2.3 Ordinary Least Squares

The first step in conducting the regression analysis was determining the transformations, if necessary, for the variables. In this particular case, each variable, if not already, was transformed to an intensive variable by dividing by either area or population. Population is an extensive variable (the same everywhere within the area), but population density is an intensive variable (varies across the region) (personal communication, Goodchild, 2004). Once the variables were appropriately transformed, the expected relationship between the independent variables and the dependent variable (number of users for each country) was stated. Then, for each a regression line was computed for each independent variable against the dependent, number of users.
Following this, a subset of variables considered to be the most relevant to describing the number of users for a country was chosen, and a regression model was developed. From this starting point, negative step regression was used to eliminate variables to develop a conservative model.

3. RESULTS

Included here is a summary of the data collected, as well as the output from both the Self-Organizing Map algorithm and the ordinary least squares regression model.

3.1 User Data

This section provides a brief overview of the users of each of the data sets. Table 3 (UNEP/GRID Arendal), Table 4 (UNEP/GRID Geneva), Table 5 (GLCC), Table 6 (GM1.0, Japan), Table 7 (GM0, Japan), and Table 8 (GM1.0, Australia) show the top five countries in terms of number of users, top 5 research sites, domains, and application categories for each of the data sets evaluated. Table 9 provides the top ten for each category for all users of all data sets combined.

3.1.1 UNEP/GRID Arendal

Table 3 provides a brief overview of the top five entries in each category for the UNEP/GRID Arendal user data. As shown in the table, Norway, the home of UNEP/GRID Arendal, has the highest number of users (19% of all Arendal users), followed by the United States. Most uses of the data downloaded were for global applications, most users were from academic institutions, and mapping was the most often given explanation for use of the data.

Table 3: Top 5 List for UNEP/GRID Arendal

	Users	#	Research	#	Domain	#	Applic.	#
1	Norway	129	Global	132	Academic	234	MAP	254
2	United States	91	Arctic	85	Personal	107	CLIM	89
3	Germany	49	no research site given	65	NGO	100	EDC	69
4	Russia	31	Europe	59	Commercial	65	blank	68
5	Canada	27	Russia	32	Government	48	POLL	44

3.1.2 UNEP/GRID Geneva

UNEP/GRID Geneva, unlike all the other data sets, does not have the most users from the same country as the data provider (Switzerland), although Switzerland does appear in the top 5. Unfortunately, not all users provided the research extent for which they downloaded the data, hence the large number without a research site given. Again, the majority of the users are from academic institutions, with the largest number of actual applications being media, press, or education (the largest number was for those with no application provided).

Table 4: Top 5 List for UNEP/GRID Geneva

	Users	#	Research	#	Domain	#	Applic.	#
1	United States	586	no site given	1794	Academic	967	no application given	1025
2	United Kingdom	193	global	61	Commercial	258	MPRED	247
3	Germany	160	Africa	24	NGO	250	RES	155
4	Japan	115	Europe	22	Government	233	EDC	77
5	Switzerland	105	South America	9	Personal	193	CLIM	65

3.1.3 Global Land Cover Characteristics (GLCC)

The GLCC data users mostly come from the United States, have global applications in mind for the data, come from academic institutions, and are mostly interested in global modelling efforts, either climate change or land use/land change. Again, many users did not provide the information regarding what application they would be using the data for.

Table 5: Top 5 List for GLCC

	Users	#	Research	#	Domain	#	Applic.	#
1	United States	447	Global	360	Academic	436	no application given	280
2	Germany	65	United States	91	Government	160	LULC	147
3	United Kingdom	59	Africa Europe	87	Commercial	148	EDC	126
4	Canada	48	North America	63	NGO	147	CLIM	90
5	Italy	43	South America	59	Personal	124	MODEL	78

3.1.4 Global Map Version 1.0, served from GSI Japan

Global Map Version 1.0 and Version 0, served from GSI Japan, both have the majority of users coming from Japan. The same for the research site: most popular is Japan. The majority of users come from personal domains, and they are interested in using the data as a reference tool.

Table 6: Top 5 List for Global Map Version 1.0, from GSI Japan

	Users	#	Research	#	Domain	#	Applic.	#
1	Japan	2650	Japan	2050	Personal	1517	REF	1351
2	United States	190	Thailand	316	Commercial	754	MPRED	597
3	Australia	171	Australia	258	Academic	553	RES	383
4	Thailand	47	Philippines	183	Government	387	REC	301
5	South Korea	28	Nepal	173	NGO	250	MAP	274

3.1.5 Global Map Version 0, served from GSI Japan

Table 7: Top 5 List for Global Map Version 0, from GSI Japan

	Users	#	Research	#	Domain	#	Applic.	#
1	Japan	678	Japan	296	Personal	404	REF	484
2	United States	62	Asia	156	Commercial	317	MPRED	158
3	South Korea	36	Global	109	Academic	190	RES MAP	121
4	Australia	33	Europe	80	Government	116	REC	85
5	Germany	30	South America	33	NGO	79	SW	44

3.1.6 Global Map Australia Version 1.0, served from GeoSciences Australia

For Global Map 1.0 Australia, the majority of the users come from Australia. All the research sites are in Australia. Again, many of the users are from the personal domain, with most using the data for reference purposes.

Table 8: Top 5 List for Global Map Australia (version 1.0)

	Users	#	Research	#	Domain	#	Applic.	#
1	Australia	2829	Australia	3511	Personal	1928	REF	1166
2	United States	173			Commercial	729	MPRED	686
3	United Kingdom	65			Government	402	EDC	282
4	Germany	59			Academic	401	URBPL	255
5	Canada	38			Media	40	TRANSP	240

3.1.7 All Data Sets Combined

Table 9 provides a brief summary for each category (which country has the most users, where research is occurring, domain, and application). In addition, the percent of the total world population for each of the top ten countries in terms of users is presented. For example, the number 1 country for users overall is Japan, with 3496 users, which is 29% of all users. Japan, however, only has 2% of the world's population. Of the top 10, all are part of what is considered the developed world, 6 of the 10 are in Europe.

Table 9: Top 10 list for all data sets combined

	Users (% of world pop)	# (%)	Research	# (%)	Domain	# (%)	Applic.	# (%)
1	Japan (2.0%)	3496 (29%)	Australia	3810 (32%)	Personal	4273 (35%)	REF	3068 (25%)
2	Australia (0.3%)	3105 (26%)	Japan	2350 (19%)	Academic	2781 (23%)	MPRED	1744 (14%)
3	United States (4.5%)	1549 (13%)	(blank)	1911 (16%)	Commercial	2271 (19%)	(blank)	1511 (13%)
4	Germany (1.3%)	384 (3%)	Global	662 (5%)	Government	1346 (11%)	MAP	852 (7%)
5	United Kingdom (1.0%)	381 (3%)	Thailand	323 (3%)	NGO	827 (7%)	EDC	685 (6%)
6	Canada (0.5%)	236 (2%)	Europe	248 (2%)	Education	265 (2%)	RES	671 (6%)
7	Switzerland (0.1%)	179 (1%)	Asia	214 (2%)	(blank)	122 (1%)	REC	591 (5%)
8	Italy (0.9%)	178 (1%)	Philippines	204 (2%)	Media	104 (1%)	URBPL	444 (4%)
9	Norway (0.1%)	174 (1%)	Nepal	177 (1%)	Military	51 (0.5%)	TRANSP	296 (2%)
10	France (0.1%)	156 (1%)	Africa	157 (1%)	Museum	22 (0.1%)	CLIM	285 (2%)

3.2 Self-Organizing Map

The output from the self-organizing map algorithm, in which the country level data set was input, is shown in Figure 1. This is the U-Matrix output (Ultsch 1999), and the darker the cell, the more distant it is from its neighbor. Therefore, those countries on the right hand side of the image are more like each other than those on the left. Australia, Japan, and the US are particularly different (the extreme darkness of the cells) from the rest of the countries.

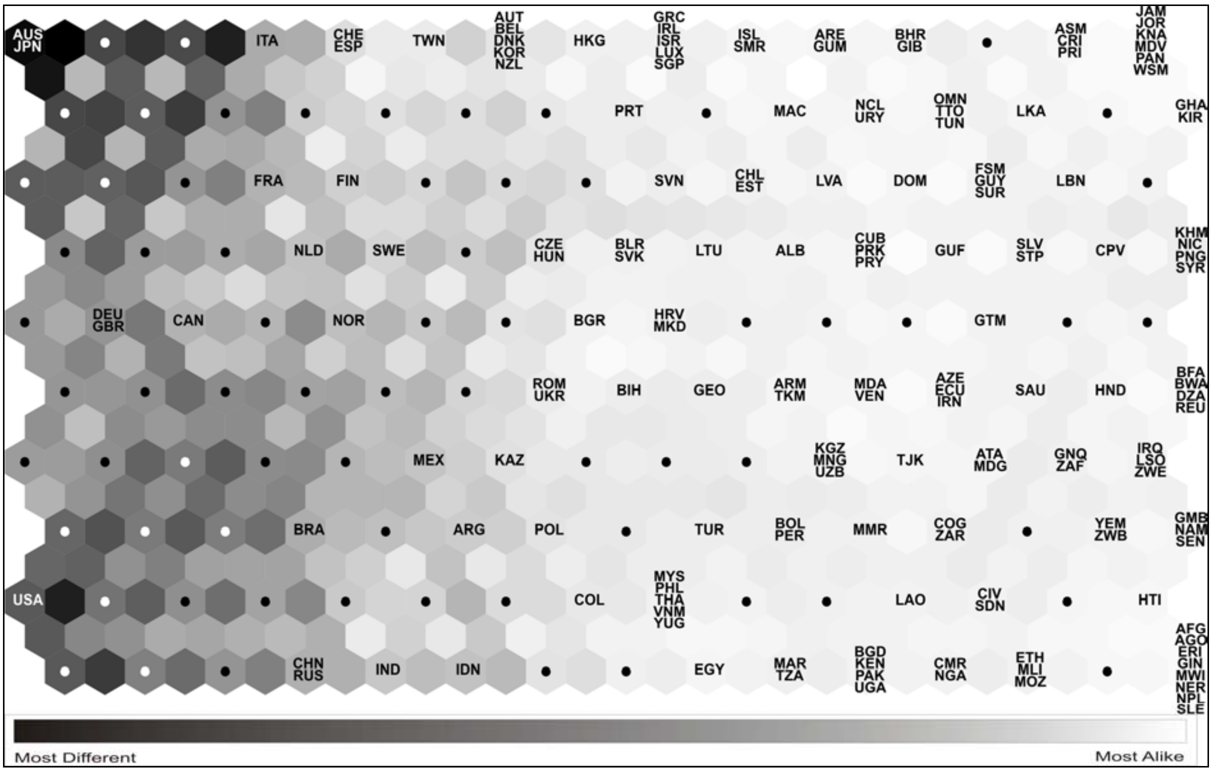

Figure 1: Self Organizing Map Output

Each of the individual component planes (each individual variable can be viewed to see where it has high and low values in the resulting SOM map) was viewed. An annotated version of the SOM output is in Figure 2, and shows where certain variables have high (red), medium (green), or low (blue) values. For example, at the center of the SOM map, where the word ‘English’ appears in blue, this area of the map shows those countries with very low values for the variable representing English language speakers.

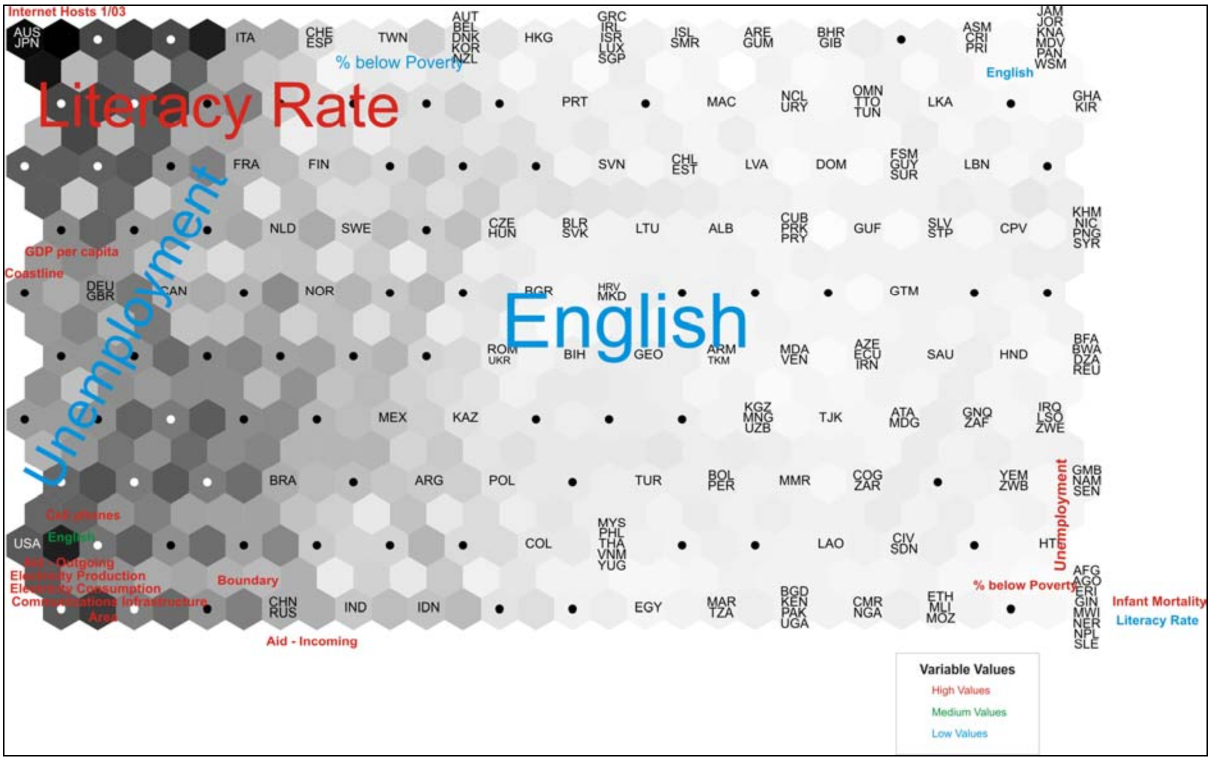

Figure 2: Annotated SOM Map Output

3.3 Ordinary Least Squares

This section provides the results of the regression models for both the individual variables (Table 10) and for the model with several variables.

Table 10: Ordinary Least Squares Regression Models for Individual Variables

Variable	Expected Relationship	Intercept	Coefficient	Significance*
Pop.Dens	Positive	2.761e-06	**3.312e-10**	__
InfMortRt	Negative	5.523e-06	**-6.607e-05**	**
LitRt	Positive	-4.735e-05	**9.731e-06**	*
GDP.pc	Positive	-7.676e-07	**4.293e-10**	***
Belpov	Negative	1.548e-06	**-2.114e-06**	***
Unempl	Negative	5.683e-06	**-1.386e-05**	.
Elect.prd.pc	Positive	3.910e-07	**9.043e-10**	***
Elect.use.pc	Positive	1.462e-07	**1.040e-09**	***
Debt.pc	Negative	-2.497e-07	**1.756e-09**	***
Aid.pc	Negative	2.942e-06	**-7.341e-10**	__
Tel.pc	Positive	-5.849e-07	**1.666-05**	***
Cell.pc	Positive	9.486e-07	**2.058e-05**	***
ISPs.pc	Positive	3.083e-06	**-8.687e-04**	__
Int.pc	Positive	-1.321e-07	**2.649e-05**	***
RR.dens	Positive	2.884e-06	**-2.854e-08**	__
Hwy.dens	Positive	2.416e-06	**5.812e-07**	__
Water.dens	Positive	2.915e-06	**-4.455e-06**	__
Airport.dens	Positive	1.981e-06	**2.123e-04**	***

Variable	Expected Relationship	Intercept	Coefficient	Significance*
Mil.pc	Positive	1.564e-06	**9.797e-09**	*
int.0103.pc	Positive	2.937e-06	**-5.054e-08**	__
int.0795.pc	Positive	1.620e-06	**1.466e-03**	***
Aid.OUT.pc	Positive	2.142e-06	**1.072e-07**	***
Aid.IN.pc	Negative	3.072e-06	**-1.288e-09**	__
English	Positive	2.257e-06	**3.228e-06**	__

Overall, the coefficients are small numbers, but the relationship (either positive or negative) is clear. The Expected Relationship is given as either positive or negative, which is compared to the sign on the coefficient. For the most part, the expected relationship was the actual relationship. However, there are some that did not end up as expected. Debt per capita was expected to be negative (more debt would result in less users); however, the relationship is positive. The number of ISPs per capita was expected to be positive (more Internet Service Providers was expected to coincide with more data users), but is, in fact, negative. Railroad density, used as a proxy for development, was expected to have a positive relationship with users (more developed would indicate more users), but the relationship is actually negative. Waterway density, also an indicator of development (shipping lanes indicating active participation in the global economy), was also expected to have a positive relationship with the number of users, but it too, has a negative relationship. The number of Internet hosts per capita in January 2003 (int.0103.pc) per capita, similar to the number of Internet service providers per capita, has a negative relationship with the number of users, despite expecting a positive relationship.

Users per capita = 3.926e-08 +
-9.743e-10 (Population Density) +
-5.635e-06 (Infant Mortality Rate) +
6.440e-07 (Literacy Rate) +
3.209e-07 (Percent Population below Poverty) +
8.843e-11 (Electricity Production per capita) +
7.341e-08 (Outgoing Aid per capita) +
1.052e-06 (English language) +
9.556e-06 (Internet hosts per capita January 2003).

In this particular model, outgoing aid per capita is significant to 0.001 while English is significant to 0.01. The adjusted R^2 value is 0.7042. Removing any other variables using negative stepwise regression results in a model with no significant variables and an adjusted R^2 value of 11%.

Here, population density and infant mortality rate have a negative relationship with the number of users per capita while the rest of the variables have a positive relationship. But overall, the number of users per capita can be described relatively well by the population density, the infant mortality rate, the literacy rate, the percent of the population below poverty, electricity production per capita, outgoing aid per capita, English language usage, and the number of Internet hosts per capita in January 2003.

4. CONCLUSIONS

Based upon the information provided here, most users download data that is served from within their home country. Not all data sets are created equal when it comes to users – some draw more from the academic/NGO research community, while others do not. For example, while Global Map was developed specifically as a response to Agenda 21, the users of the Global Map data are far and above people just looking at the data set to see what it is about. While the GLCC data, on the contrary, were developed for and are being used by the research community, particularly modelers.

Global Map data have an additional hurdle – while it is global in conception, it is not global in production. Each country produces their own contribution, which is then made available at the country level from the ISCGM servers in Tsukuba, Japan. This is contrary to the GLCC data, which is a global data set, the smallest sections of which can be accessed are at the continent level. If the ISCGM wants to increase the visibility of Global Map outside of Japan, then a concerted marketing effort needs to be made around the world. However, the ISCGM would most likely balk at this idea – marketing takes lots of time and money. The Global Map project participants are doing this on a voluntary basis, and there is little to no money available for a large-scale worldwide marketing campaign.

However, it is possible, since it was noted that users tend to find data sets within their own country's data servers, that each country could serve their Global Map contribution from their own government agency servers. For example, the USGS could provide the U.S. contribution to Global Map from within the USGS website, in addition to the provision of the same data from the ISCGM website, based in Tsukuba, Japan.

Another issue that Global Map must take a look at is the fact that many of their users are what are considered "Personal" in this context. Since Global Map was developed specifically as a response to the call for data in Agenda 21 from the UNCED meeting, it would be expected that there would be a significant presence of NGO users (United Nations employees fall within this category). However, this is not the case, particularly with the overwhelmingly large number of personal users. Global Map needs to clearly define who it's targeted user audience is, and then reach out to that audience to ensure it knows the data exist, how they can access the data, and what they can do with the data.

Data stored in an electronic medium of any kind need a steady electricity supply as well as a steady telecommunications infrastructure for them to be accessed and used effectively. Civil war and upheaval have a tendency to often limit the steady supply of electricity as well as other traditionally taken for granted infrastructure, such as potable water supply. Not having these types of basic infrastructure available would tend to not just limit a person's use of data, but their interest in learning how to find it and use it, particularly when they wonder whether they are safe or where their next meal will come from.

While the solutions to these problems are nontrivial, and are long term issues without a solution in the near term, they need to be kept in the mind of those organizations trying to increase the use of their data sets. Just simply placing the data online, or limiting the announcement of the data availability to a single country or region is not enough.

A ‘digital divide’ between those countries with computing and communications infrastructure and those without is clearly visible using the SOM (Figure 3). This clearly indicates that work at the political, legal, and infrastructural level must be considered and initiated if the user base of global cartographic data sets is to be increased, particularly for those outside of the major data consumers – the U.S., Australia, Europe, and Japan.

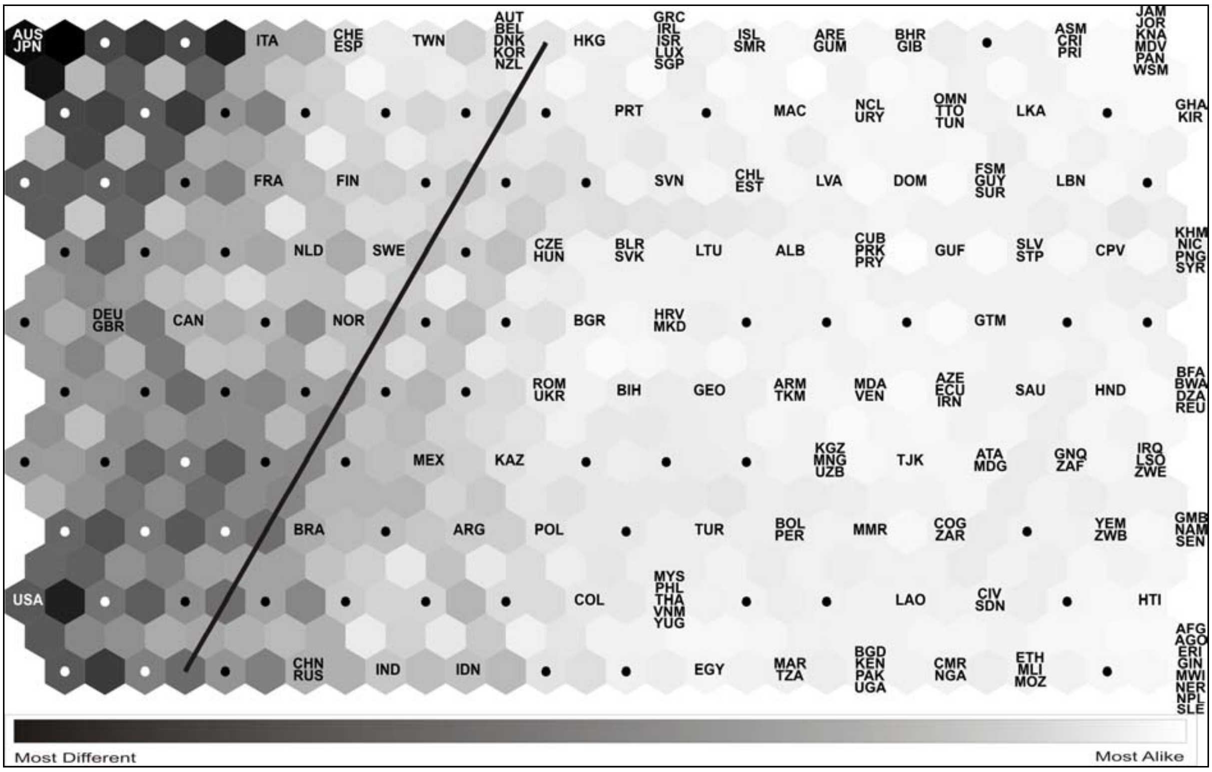

Figure 3: Location of the Digital Divide

In conclusion, one point becomes clear: All data are local. Users are more interested in locating and utilizing data that reference geographical locations near them. In addition, several other points are made:

- Education is a necessary prerequisite for the appropriate use of global environmental data sets. Without education users are not able to fully benefit from the data.
- Identification of current non-user communities that may benefit from use of these data should be done.
- Electrical and communications infrastructure are necessary for accessing and utilizing the data, and are not necessarily well established every where yet.
- Data provision should be made in other languages besides English.

REFERENCES

Balakrishnan, N. (2001). "Information and communication technologies and the digital divide in the Third World countries." Current Science **81**(8): 966-976.

Central Intelligence Agency (2002). The World Factbook 2002, Central Intelligence Agency. **2002**.

Dasgupta, S., S. Lall, et al. (2001). Policy Reform, Economic Growth, and the Digital Divide: An Econometric Analysis, World Bank**:** 18.

Deboeck, G. (1998). Picking Mutual Funds with Self-Organizing Maps. Visual Explorations in Finance with Self-Organizing Maps. G. Deboeck and T. Kohonen. London, Springer**:** 39-58.

Department of Commerce (1999). Falling Through the Net: Defining the Digital Divide. Washington, DC, National Telecommunications and Information Association. **2003**.

Hadley, B. C., J. E. Estes, et al. (2000). European and Asian Satellite Remote Sensing Applications: A Literature Review and Analysis. Santa Barbara, University of California**:** 172.

James, J. (2003). Bridging the Global Digital Divide. Cheltenham, UK, Edward Elgar.

Kangas, J. and T. Kohonen (1996). "Developments and applications of the self-organizing map and related algorithms." Mathematics and Computers in Simulation **41**: 3-12.

Kohonen, T. (2001). Self-Organizing Maps. Berlin, Springer.

Kohonen, T., S. Kaski, et al. (2000). "Self Organization of a Massive Document Collection." IEEE Transactions on Neural Networks **11**(3): 574-585.

Kohonen, T. and P. Somervuo (2002). "How to make large self-organizing maps for nonvectorial data." Neural Networks **15**: 945-952.

Organisation for Economic Co-Operation and Development (2002). Geographical Distribution of Financial Flows to Aid Recipients, 1996-2000: Disbursements, Commitments, Country Indicators. Paris, OECD Publications.

Skupin, A. (2003). A Novel Map Projection Using an Artificial Neural Network. 21st International Cartographic Conference (ICC), 'Cartographic Renaissance', Durban, South Africa, International Cartographic Association.

Ultsch, A. (1999). Data Mining and Knowledge Discovery with Emergent Self-Organizing Feature Maps for Multivariate Time Series. Kohonen Maps. E. Oja and S. Kaski. Amsterdam, Elsevier Science**:** 33-45.

Vesanto, J. (1999). "Som-based data visualization methods." Intelligent Data Analysis **3**: 111-126.

BIOGRAPHICAL NOTES

Ms. Kline received her PhD from the Department of Geography at the University of California, Santa Barbara. She has been involved with the activities of the International Steering Committee for Global Mapping (ISCGM) since the mid-1990s, and is currently in residence as a Japan Society for the Promotion of Science (JSPS) post-doctoral fellow at the Geographical Survey Institute and Secretariat of the ISCGM.

CONTACTS

Karen D. KLINE, PhD
JSPS Post-Doctoral Fellow
Secretariat of the ISCGM
Geographical Survey Institute
Kitasato-1
Tsukuba-shi
Ibaraki-ken, 305-0811
JAPAN
Tel. +81 298 64 6910
Fax + 81 298 64 1804
Email: kline@gsi.go.jp
Web site: http://www.iscgm.org