

## Global Environmental Data Sets and the Digital Divide

*Karen D. Kline, PhD*  
Assistant Secretary General  
International Steering Committee for Global Mapping  
JSPS Research Fellow  
Geographical Survey Institute, Japan

GSDI 8/FIG Working Week  
20 April 2005

## Questions

- ▶ *Enormous expenditures on data collection and development*
  - ▶ *Are these products really useful?*
  - ▶ *For what purposes are they being used?*
  - ▶ *Are there applications that have not been identified that could be marketed?*
- ▶ *Data distribution policies are currently being written for various products*
  - ▶ *Commercial use argument*
  - ▶ *COFUR vs. full cost recovery*

## Goals

- ▶ *Improved customer service*
  - ▶ *Better understanding of the users*
    - ▶ *Distribution mechanisms*
    - ▶ *Updates to the data*
  - ▶ *Improved provider-customer relationship*
- ▶ *More understanding of the user when drafting data distribution policies*

## Data Analyzed

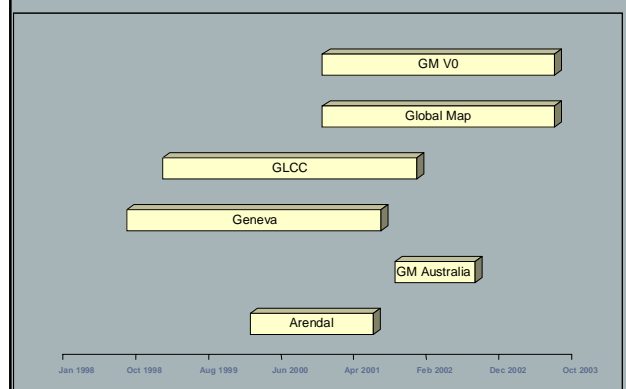
- ▶ *Global Map Version 0 (Tsukuba, Japan)*
- ▶ *Global Map Version 1 (Tsukuba, Japan)*
- ▶ *Global Map Australia (Canberra, Australia)*
- ▶ *Global Land Cover Characteristics (GLCC) (Sioux Falls, South Dakota)*
- ▶ *UNEP/GRID-Arendal (Arendal, Norway)*
- ▶ *UNEP/GRID-Geneva (Geneva, Switzerland)*

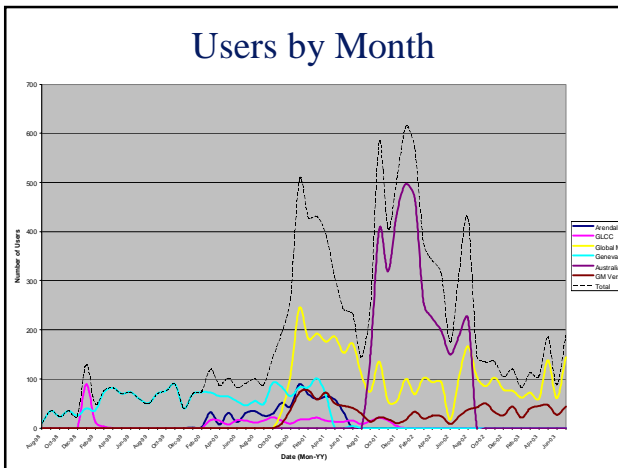
## Why these data sets?

- ▶ *User information collected via the internet when data set downloaded*
- ▶ *Organizations released user information*
- ▶ *All are based on different continents*
- ▶ *All created for different purposes*

UNEP/GRID | Global Map | GLCC

## Start and End Dates for All Data Sets



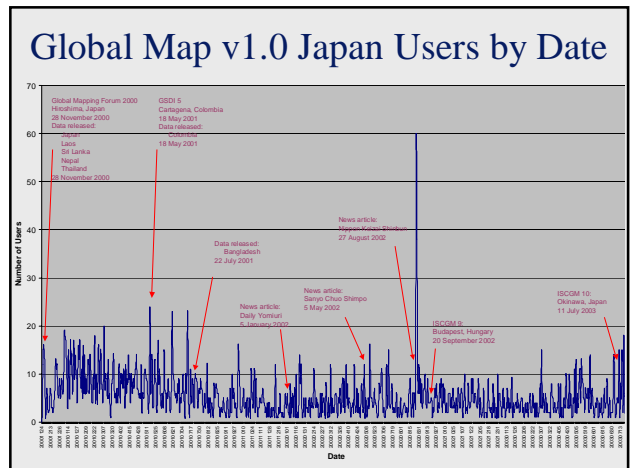


### Global Map Japan User Data

- ▲ All users must register
- ▲ 2 files:
  - ▲ Registration information
  - ▲ Downloaded files
- ▲ Merge the two, remove duplicates
- ▲ Review and code for applications
- ▲ Corrections/Interpretations
  - ▲ Japan/Jamaica
  - ▲ Japanese language
- ▲ Data Sets:
  - ▲ Global Map Version 0
  - ▲ Global Map Version 1.0 (all released countries)

### Summary of GM Version 1.0 Japan Users

- ▲ 3603 Users (29% of all users)
- ▲ November 2000-July 2003
- ▲ Main source of users: Japan (74%)
- ▲ Main research location: Japan (57%)
- ▲ Domain: Personal (42%)
- ▲ Application: Reference (37%)



### Summary of GM Version 0 Users

- ▲ 1158 Users (10% of all users)
- ▲ November 2000-July 2003
- ▲ Main source of users: Japan (59%)
- ▲ Main research location: Japan (26%)
- ▲ Domain: Personal (35%)
- ▲ Application: Reference (42%)

### Global Map Australia User Data

- ▲ All users must register
- ▲ Certain information removed by Australian Government
- ▲ Review and code for applications
- ▲ Data Sets:
  - ▲ Global Map Version 1.0 (Australia only)

## Global Map Australia User Data

- ▲ 3511 Users (29% of all users)
- ▲ September 2001-August 2002
- ▲ Main source of users: Australia (81%)
- ▲ Main research location: Australia (100%)
- ▲ Domain: Personal (55%)
- ▲ Application: Reference (33%)

## Global Land Cover Characteristics

- ▲ User information form filled out prior to downloading
  - ▲ sent via email to USGS/EDC staff
  - ▲ registration information entered into excel file
- ▲ Date information not included with all entries in excel file
- ▲ Reviewed and coded for applications

## Summary of GLCC User Data

- ▲ 1108 Users (9% of all users)
- ▲ February 10, 1999 - December 17, 2001
- ▲ Main source of users: United States (40%)
- ▲ Main research location: Global (32%)
- ▲ Domain: Academic (39%)
- ▲ Application: no application given (25%)

## UNEP/GRID Overview

- ▲ United Nations Environment Programme, Global Resource Information Database
- ▲ Established 1995
- ▲ Regional nodes
  - ▲ Different regional emphases
  - ▲ Different scientific emphases
- ▲ GRID objective:
  - ▲ Promote exchange and availability of global and regional environmental data

## UNEP/GRID Arendal User Data

- ▲ Requests tracked by entering into database by UNEP/GRID Arendal staff
- ▲ General manual (how-to) requests ignored
- ▲ Some entries written in Norwegian

## Summary of UNEP/GRID Arendal User Data

- ▲ 671 Users (6% of all users)
- ▲ January 2000 - June 2001
- ▲ Main source of users: Norway (19%)
- ▲ Main research location: Global (20%)
- ▲ Domain: Academic (35%)
- ▲ Application: Mapping (38%)

## UNEP/GRID Geneva User Data

- ▲ Users submitted registration information prior to downloading data
- ▲ User logs provided in text format
- ▲ Not all fields required
- ▲ Each entry reviewed individually

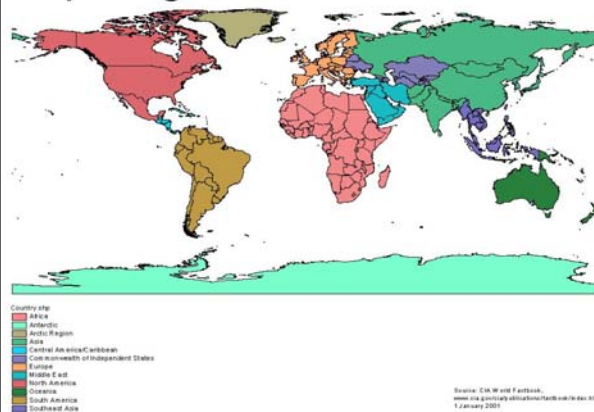
## Summary of UNEP/GRID Geneva User Data

- ▲ 2011 Users (17% of all users)
- ▲ August 1998 - July 2001
- ▲ Main source of users: United States (29%)
- ▲ Main research location: no site given (89%)
- ▲ Domain: Academic (48%)
- ▲ Application: no application given (51%)

## Data Limitations

- ▲ Information collected was not consistent
- ▲ English not the first language of many users
  - ▲ What they meant to write may not be what I understood
- ▲ No quality control in the collection process
- ▲ Information either clearly decipherable or must be inferred by looking at email addresses, physical addresses, other provided information

## Map Regions



## Domains

- ▲ Academic
- ▲ Commercial
- ▲ Education
- ▲ Government
- ▲ Media
- ▲ Military
- ▲ Museum
- ▲ NGO
- ▲ Personal

## Application Categories

- |  |                                 |
|--|---------------------------------|
| 1. Agriculture                           | 11. Marine                      |
| 2. Air Quality                           | 12. Media, press and education  |
| 3. Economic development and conservation | 13. Public health               |
| 4. Emergency management                  | 14. Rangeland                   |
| 5. Fisheries                             | 15. Recreation and tourism      |
| 6. Forestry                              | 16. Transportation              |
| 7. Geology                               | 17. Urban and regional planning |
| 8. Information and intelligence          | 18. Water quality               |
| 9. Land use and land cover               | 19. Water resources             |
| 10. Mapping, charting and geodesy        | 20. Weather and climate         |

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

## Derived Data Sets

### Users

- ▲ One entry per user
  - ▲ Data set
  - ▲ Date
  - ▲ Location of user
  - ▲ Domain of user
  - ▲ Location of research
  - ▲ Application

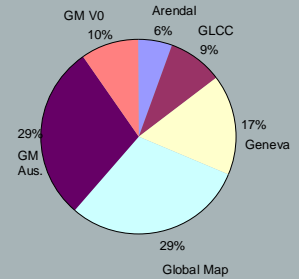
### Countries

- ▲ One entry per country
  - ▲ Number of users
  - ▲ Number of research sites
  - ▲ Socioeconomic data
  - ▲ Aid flow data
  - ▲ Communications infrastructure data

## General Statistics

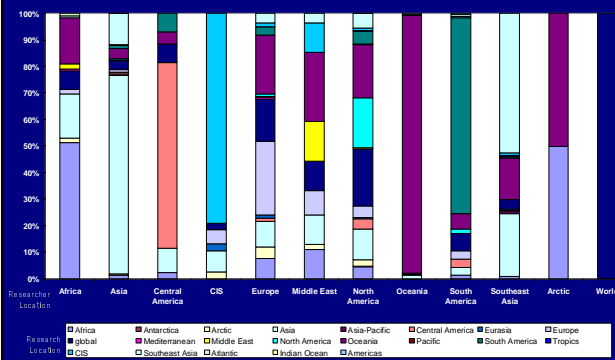
### Total Number of Users:

- ▲ UNEP/GRID Arendal 671
- ▲ GLCC 1108
- ▲ Global Map Japan 3603
- ▲ UNEP/GRID Geneva 2011
- ▲ Global Map Australia 3511
- ▲ Global Map Version 0 1158
- ▲ TOTAL 12062



## Research v. Researcher Location

All Data Sets



## Top Applications by Data Set

Arendal	GLCC	GM Japan	Geneva	GM Australia	GM V0
MAP	(blank)	REF	(blank)	REF	REF
CLIM	LULC	MPRED	MPRED	MPRED	MPRED
EDC	EDC	RES	RES	EDC	MAP
(blank)	CLIM	REC	EDC	URBPL	RES
POLL	MODEL	MAP	CLIM	TRANSP	REC

## Top Research Locations by Data Set

Arendal	GLCC	Geneva	Global Map Japan	Global Map Australia	Global Map V0
Global	Global	(blank)	Japan	Australia	Japan
Arctic	United States	Global	Thailand	---	Asia
(blank)	Europe	Africa	Australia	----	Global
Europe	Africa	Europe	Philippines	----	Europe
Russia	North America	South America	Nepal	---	Southeast Asia

## Top Domains by Data Set

Arendal	GLCC	Geneva	Global Map Japan	Global Map Australia	Global Map V0
Academic	Academic	Academic	Personal	Personal	Personal
Personal	Government	Commercial	Commercial	Commercial	Commercial
NGO	Commercial	NGO	Academic	Government	Academic
Commercial	NGO	Government	Government	Academic	Government
Government	Personal	Personal	NGO	Media	NGO

## Data Observations

- ▲ Majority of users from same country as provider
- ▲ Most users conduct research in the same country they are located in
- ▲ Internet connectivity necessary for data use

## Self Organizing Map

- ▲ An unsupervised neural net algorithm for visualization of multivariate data
- ▲ Objects near each other in multivariate space are displayed near each other in the SOM output map

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

## Data Input to SOM

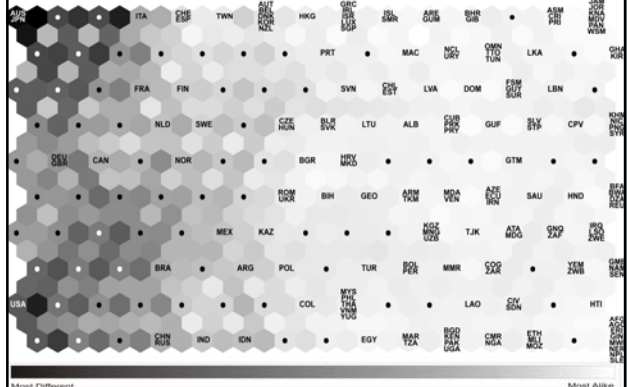
- ▲ Country level data based on:
  - ▲ The information in the user database
  - ▲ Internet host information
  - ▲ Socioeconomic data
  - ▲ Financial aid flow data

Sources:

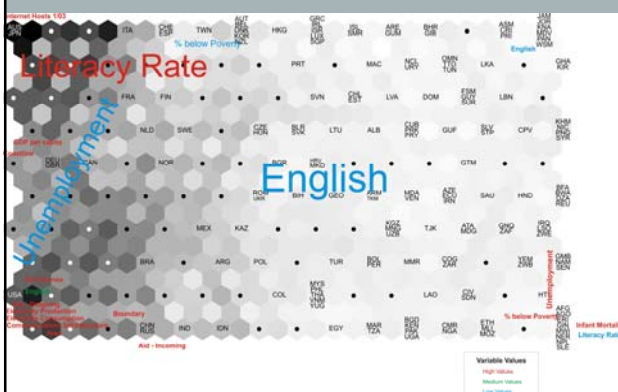
•Central Intelligence Agency (2002). *The World Factbook 2002*. Central Intelligence Agency.  
 •Organisation for Economic Co-Operation and Development (2002). *Geographical Distribution of Financial Flows to Aid Recipients, 1996-2000*. Disbursements, Commitments, Country Indicators. Paris, OECD Publications.  
 •Internet Software Consortium, www.isc.org

## SOM

### All Countries w/Positive Users & Research

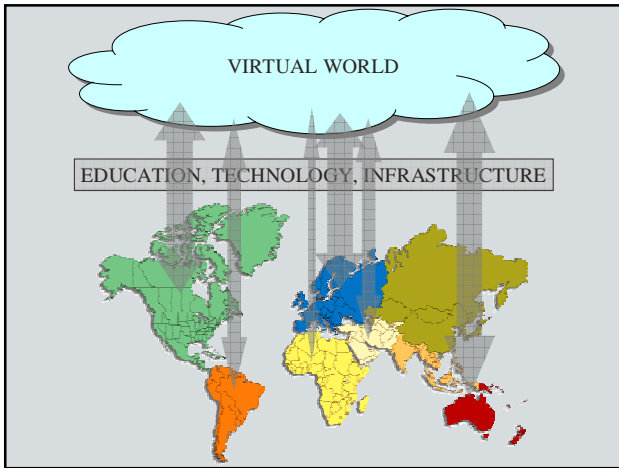
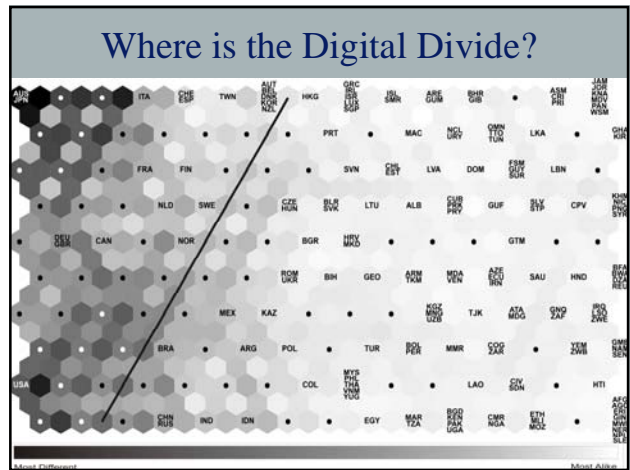
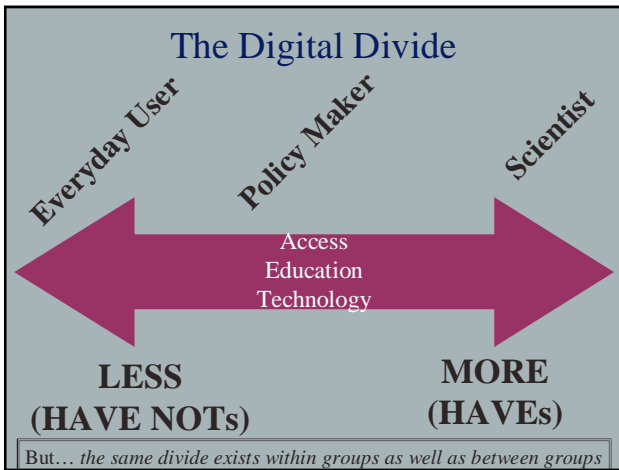


## SOM – Socioeconomic Variables



## Digital Divide

- ▲ The gap between those who have and those who do not have...
  - ▲ Computing technology
  - ▲ Purchasing ability
  - ▲ Education
  - ▲ Technological skills
- ▲ U.S.:
  - ▲ Rural/urban focus
  - ▲ Telephone connection (ITU)
- ▲ Global:
  - ▲ Developed/less developed focus (UNDP)



### You can give me all the data in the world, but if I don't have...

- ▲ The education to convert the data to useful information AND
- ▲ A computer with enough memory and processing power AND
- ▲ The ability to read the media on which the data were provided AND
- ▲ The time to do so OR
- ▲ The money to pay someone to do it for me

Then all the data in the world is useless.

### Acknowledgements

Tom Loveland, Jess Brown, Stephen Howard (EROS Data Center)  
 Tim Foresman, Mick Wilson, Lawrence Hislop, Ron Witt, Hy Dao, Lorant Czarán (UNEP)  
 Minoru Akiyama, Hiroshi Une, Hiroshi Masaharu, Eiji Murakami, Hidenori Fujimura, Hiromichi Maruyama (Global Map/GSI)  
 Peter Holland (GeoSciences Australia)

John Kelmelis, USGS; Peter Holland, GeoSciences Australia; Claude Luzet, EuroGeographics; Santiago Borrero-Mutis, PAIGH

Keith C. Clarke, Michael F. Goodchild, Leal A.K. Mertes, K. Eric Anderson

Support provided by:

- ▲ US DOI 00HQAG0018 Global Mapping Project
- ▲ NASA NAG5-10457 Remote Sensing Information Science Research
- ▲ Japan Society for the Promotion of Science (JSPS) Post-Doctoral Fellowship for North American and European Researchers (Short-Term)