

GGD-BROWSE: BRIDGING THE GAP BETWEEN DATA DESCRIPTIONS AND DATA

Christopher Haggerty, Claire Hanson

*National Snow and Ice Data Center/World Data Center-A for Glaciology
Boulder, Colorado, 80309-0449 USA
e: mail: haggerty@kryos.colorado.edu*

Abstract

The World Data Center-A for Glaciology, in cooperation with the International Permafrost Association (IPA), has developed a data discovery tool for the frozen ground scientific community. The Global Geocryological Database (GGD) contains data and descriptions for world-wide frozen ground observations. The GGD-Browse tool provides data search and retrieval capability based on extensive metadata in the GGD. This browse tool is designed primarily for use through the World Wide Web (WWW), but it can easily be adapted for use on portable media such as CD-ROMs. Search capabilities based on commonly used parameters such as principal investigator, region, geographic area, and topic are included within the system.

The number of data sets, modeling packages, and other information contained within the GGD continues to grow, and GGD-Browse is designed to grow as well. By combining a relational database with the WWW, extension of the interface is designed to be a simple process.

Introduction

The advent of new electronic media in the 1980's began to provide scientists with a plethora of shared data. Data used in scientific studies could be shared as easily as the results of such studies had been shared in the past. Digital storage media and the microcomputer began to complement journals and conferences as the means of scientific exchange. The current technology of the World Wide Web and CD-ROMs allows the global scientific community to obtain gigabytes of data, but the amount of data accessible is not always matched by ease of data access. The raw nature and sheer mass of available data can overwhelm the user.

The design of data collections, and the management of these collections, can provide a level of structure needed for scientific usage. The design of data and information has come to be called information architecture. The design process of the Global Geocryological Database Browse (GGD-Browse) system was governed by the principles of information architecture to create a system which allows the user access to data in a clear and rapid fashion.

Information architecture

To begin to understand the design process we must apply a model of learning and information integration that describes how people take in sensory data and

intellectually integrate these data. This process is defined as the "continuum of understanding" (Shedroff, 1994). The GGD-Browse design process utilizes a series of organizational tools to build a user interface from new data to link to a user's experiential knowledge.

Data themselves are the beginning stage in a process of understanding. This process (Figure 1) can be seen as a progression from Data to Information, thence to Knowledge, and finally to Wisdom. The concept of the progression from raw measurements through a research process leading to a published result is familiar to most of the scientific community.

The transition from Data to Information is facilitated by the organization of data in a manner that communicates the relationships within the data themselves. The metadata furnishes tools to the data product designer to organize the data into categories with which users are familiar (Wurman, 1996) and that define the necessary structure for the transition from Data to Information. Using these methods of organization provides a connection to universal knowledge and thus builds on the user's experience.

GGD-Browse

The GGD contains a series of metadata records, site descriptions, and other documentation about frozen ground and permafrost observations. By combining the raw observations and metadata into a structured envi-

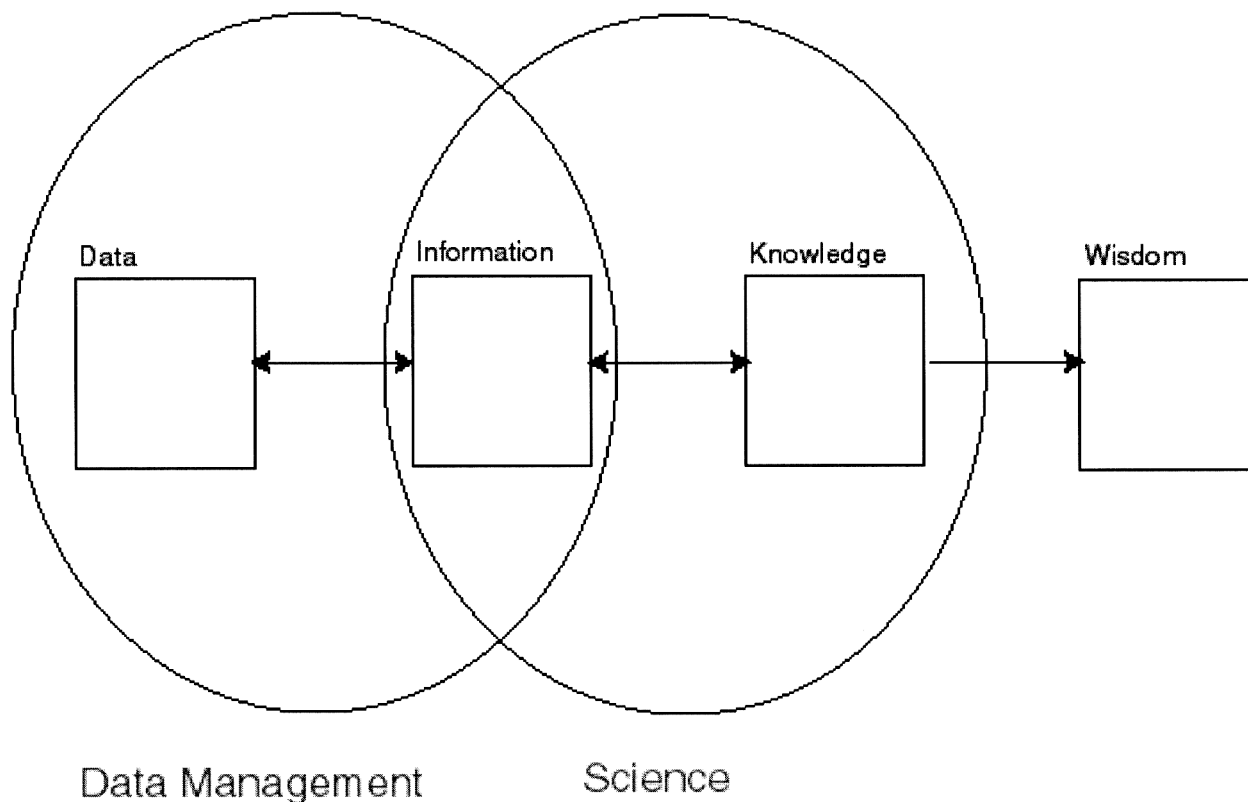


Figure 1: The "continuum of understanding" (Shedroff, 1994) illustrating the linking of data management and scientific research activities leading from data and information to knowledge and wisdom.

ronment, the GGD is designed to provide Data and Information to the permafrost community. The GGD-Browse system (Figure 2) uses this Information and its underlying structure to provide a search interface that will speed the use of these Data in ongoing research.

The integration of Data and Information into Knowledge is the process of science that individual researchers undertake. Within the scientific community the continuum of understanding flows both ways. For example, through publishing research results an investigator contributes Knowledge to the community, while the same investigator may contribute Data to the GGD. The goal of GGD-Browse is to provide an environment where researchers can find Data related to published Knowledge as well as providing a structure where Data and Information are developed into new Knowledge.

Based on the assumption that all Information within the GGD-Browse system must be a conduit in two directions, the system was designed with organizing categories of Location, Author/Principal Investigator, Time, and Keyword. By using a search mechanism based on Author/Principal Investigator, GGD-Browse will bridge the gap between Knowledge (from publications) and the underlying Data. The structure of Location, Time, and Keyword allows user searching based on the common experience of the permafrost community. The organizing categories probably will

evolve with time as feedback from the user community is incorporated.

The visual nature of combining graphics with text allows the search tools to be structured in a easy-to-use "point and click" manner. This ability lets users point at locations on a map or select from a list of available Keywords. By giving the user choices and making maps interactive, potential linguistic barriers can be partially overcome, speeding a user toward the Data.

Using the site descriptions and metadata required by the GGD allows all data to be searched using GGD-Browse even if the Data themselves are not archived at NSIDC. Using Web-based technology GGD-Browse can search through the Information about Data that are stored world-wide.

CD-Rom and the world wide web

The Data and Information Working Group of the International Permafrost Association is coordinating a major state-of-the-art review of frozen ground-related data and information culminating in an extended session at the Seventh International Conference on Permafrost, June 1988 at Yellowknife, N.W.T., Canada (Hanson et al., 1997). The review session will coincide with the launch of a significant CD-ROM provided to all delegates and intended for a broad audience inclu-

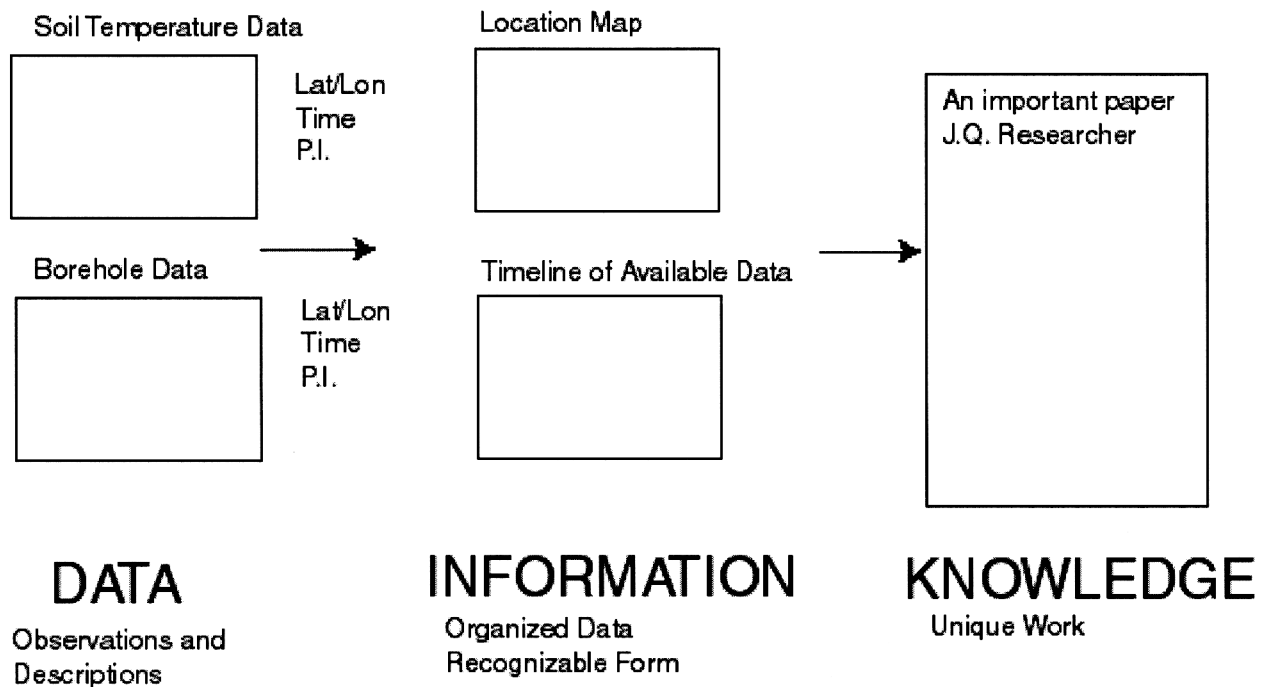


Figure 2: GGD-Browse system design allows searching for data using information (place, time, type of data) or from "knowledge", such as the author/investigator's published paper.

ding the scientific and engineering communities, educators and policy makers.

The Circumpolar Active-Layer Permafrost System (CAPS) CD-ROM will contain the digital version of the IPA Northern Hemisphere permafrost map; a bibliography of permafrost literature spanning 1978 - 1996; a glossary of frozen ground-related terms in twelve languages; a bibliography and index of more than 600 Russian permafrost maps; and selected data sets of active layer and borehole temperature profiles, representative of important time series from many of the member countries of the IPA.

To meet the challenge of providing a user interface for the GGD and the CAPS CD-ROM, GGD-Browse has been designed to be a portable system. By basing GGD-Browse on the use of hypertext (i.e. HTML) to provide structure along with textual and visual materials, the problem of supporting many different kinds of computer platforms is minimized. The popularity of the World Wide Web (WWW) has encouraged the development of several freely available hypertext browsers for a large variety of platforms. Hypertext browsers can operate on documents both across the WWW as well as on local storage media such as CD-ROM, allowing the use of a single design for the GGD and the CAPS CD-ROM.

The variety of Data within the permafrost community is mirrored by the spectrum of Data on the CAPS CD-ROM. Thus the CAPS CD-ROM provides a first introduction to the GGD and GGD-Browse. As the amount of Data within the GGD continues to grow, the hyper-

text-based interface allows updates to be provided to users through the World Wide Web. By using the same design on both the CAPS CD-ROM and the WWW based GGD-Browse, a user need learn only one such methodology. Additionally, new technologies such as Java data display tools can be integrated into the current product by using the hypertext ability to link to new document formats.

Conclusion

By using information architecture principles in designing the interface for the Global Geocryological Database (use of underlying structure to provide both Data and Information to the user), NSIDC has produced a data discovery tool that will allow scientists around the world to access Data and Information about frozen ground. The structure upon which GGD-Browse has been designed allows growth in both the number and types of Data contained within the system.

References

- Hanson, C.S., Branson, J., Clark, M.J. and J. Brown.** (1997). Circumpolar Active-Layer Permafrost System (CAPS): *A global geocryological database CD-ROM*. In Lewkowicz, A.G. and Kokelj, S. (eds.), Program and Abstracts 27th Arctic Workshop, Ottawa, 27 February - 2 March 1997, p. 104.
- Wurman, R. S. and Bradford, P.** (1996). *Information Architects*. Zurich, Graphis Press.
- Shedroff, N.** (1994). Information Interaction Design: A Unified Field Theory of Design.
[http:// www.vivid.com/form/unified/unified.html](http://www.vivid.com/form/unified/unified.html)