

# DISCRETE GLOBAL GRIDS FOR DIGITAL EARTH

Michael F. Goodchild<sup>1</sup>

## Abstract

Digital Earth is being promoted as a framework into which "we can embed vast quantities of information about the planet" (U.S. Vice President Gore). It provides an organizational metaphor for information about the Earth by allowing users to search over a virtual rendering of the planet; a call for a framework that can accommodate all kinds of information, including data, computational models of processes, text, and images; and a high-end visualization system allowing rapid change of resolution from global to local. These three interpretations all motivate a search for novel hierarchical data structures, but present distinct problems. The paper reviews the various interpretations of Digital Earth, examines progress in various arenas, and develops a series of principles for the design of global gridding schemes.

## Concepts of Digital Earth

Information in digital form can often be regarded as a *representation* of some aspect of reality, and this perspective is particularly useful in connection with geographic phenomena, that is, phenomena located on or near the surface of the Earth. Geographic representations are normally confined to spatial resolutions of centimeters or coarser, and extents that range from neighborhoods up to the entire Earth. They include maps, images of Earth from space, textual descriptions of places, spoken records, and digital databases using raster and vector structures.

In recent years there has been increasing interest in holistic approaches to representation. Projects with names like *Virtual Human* or *Virtual Los Angeles* aim to create digital representations of entire systems, and to provide the tools that allow users to explore them interactively, often in immersive environments. A user of Virtual Los Angeles ([www.gsaup.ucla.edu/bill/LA.html](http://www.gsaup.ucla.edu/bill/LA.html)), for example, will be able to explore various aspects of the city through immersion in an environment that appears to have many of the characteristics of the real city. Of course the underlying representations are limited to certain ranges of resolution, certain easily captured aspects such as built form, and certain intervals of time. But given adequate data and sufficiently powerful computing environments, these projects offer enormous value to educators, prospective tourists and entrepreneurs, city managers, and many others. Virtual anatomy ([www.vis.colostate.edu/library/gva/gva.html](http://www.vis.colostate.edu/library/gva/gva.html)) could provide the opportunity to work with a representation of a generic human body, and thus could offer medical schools great advantages in training students under carefully controlled and relatively risk-free conditions. Virtual humans and virtual Earth share many characteristics: the need for unambiguous methods for referring to locations within the frame of the whole system; the potential for representation of dynamic processes as well as static form; the need for advanced visualization systems; and similar motivations.

---

<sup>1</sup> National Center for Geographic Information and Analysis, and Department of Geography, University of California, Santa Barbara, CA 93106-4060, USA. Phone +1 805 893 8049. FAX +1 805 893 3146. Email [good@ncgia.ucsb.edu](mailto:good@ncgia.ucsb.edu).

From this perspective, a virtual or *Digital Earth* would be a unified repository and source for all that is known about the Earth system. It could provide a "one-stop shop" for geospatial data, overcoming many of the problems that are now associated with geospatial data acquisition (data scattered among different servers, in different formats, with different access protocols; redundancy in data production; etc.). It could also contain knowledge about the Earth that is stored in books, journals, reports, and other text-dominated media. And perhaps most importantly, it could contain digital representations of the processes that operate on the Earth's surface and near-surface—the numerical simulation models that represent what we know about how the Earth operates, and how it will change in the future.

It is widely accepted that the term *Digital Earth* (DE) originates with the published version of a speech of the U.S. Vice President, Al Gore. In it, he describes an immersive environment that would allow its users to explore and learn about the Earth and its human and physical environments (the full text is at [www2.nas.edu/besr/238a.html](http://www2.nas.edu/besr/238a.html); a summary was delivered in Los Angeles in January 1998):

“Imagine, for example, a young child going to a Digital Earth exhibit at a local museum. After donning a head-mounted display, she sees Earth as it appears from space. Using a data glove, she zooms in, using higher and higher levels of resolution, to see continents, then regions, countries, cities, and finally individual houses, trees, and other natural and man-made objects. Having found an area of the planet she is interested in exploring, she takes the equivalent of a ‘magic carpet ride’ through a 3-D visualization of the terrain. Of course, terrain is only one of the numerous kinds of data with which she can interact. Using the system’s voice recognition capabilities, she is able to request information on land cover, distribution of plant and animal species, real-time weather, roads, political boundaries, and population. She can also visualize the environmental information that she and other students all over the world have collected as part of the GLOBE project. This information can be seamlessly fused with the digital map or terrain data. She can get more information on many of the objects she sees by using her data glove to click on a hyperlink. To prepare for her family’s vacation to Yellowstone National Park, for example, she plans the perfect hike to the geysers, bison, and bighorn sheep that she has just read about. In fact, she can follow the trail visually from start to finish before she ever leaves the museum in her hometown.

She is not limited to moving through space, but can also travel through time. After taking a virtual field-trip to Paris to visit the Louvre, she moves backward in time to learn about French history, perusing digitized maps overlaid on the surface of the Digital Earth, newsreel footage, oral history, newspapers and other primary sources. She sends some of this information to her personal e-mail address to study later. The time-line, which stretches off in the distance, can be set for days, years, centuries, or even geological epochs, for those occasions when she wants to learn more about dinosaurs.”

Several principles and challenging ideas underlie this piece of technological fantasy. First, the immersive environment provides a very rich form of communication between the information store and the learner, unimpeded by the constraints of a single medium,

and not limited to the visual channel or to the traditional and narrow concept of a paper map as we normally understand it. Many of the constraints of traditional mapping disappear in the DE environment, including the need to project the Earth onto a flat surface, and the limitation of an unmodifiable and therefore static display. Second, the vision mixes types of data that are readily communicated by *rendering* into something resembling their true appearance, such as topography and land cover, with other types that will have to be communicated symbolically. This second type includes information on population, health, or environmental quality. Cartographers are familiar with the problems of mixing these two types through their experience with symbolic enhancement of orthographic images. Other information mentioned in the speech is geographic only in the sense of having a footprint; the contents of newspapers and oral histories will have to be represented iconically, and their contents communicated in some appropriate way, since they are not geospatial and therefore cannot be mapped onto the Earth's surface.

More fundamentally, perhaps, DE embodies a novel metaphor for the organization of digital information and construction of user interfaces. The current generation of computer operating systems, such as Windows 98, makes use of the metaphor of the desktop, with its clipboards, filing cabinets, and briefcases, because this is the environment most familiar to office workers. This tradition goes back to work at the Xerox PARC laboratories in the 1960s, but came to dominate Microsoft operating systems only in the late 1980s with Windows. Yet the office is not a natural environment for thinking and learning about the surface of the Earth, and *office* is not the first thing that comes to mind when we think of Columbus, or von Humboldt. Since all such information relates to some geographic location, it would be far more effective to use the Earth's surface itself as the organizing metaphor. For example, rather than look in a filing cabinet under Z, someone interested in Zimbabwe would find it much easier conceptually to reposition a digital globe to the right part of Africa (or to look up Zimbabwe in a digital rendering of the back-of-the-atlas gazetteer, and see the globe repositioned automatically). DE replaces the office with the Earth as the dominant user interface metaphor. In that sense it offers a significant contribution to the growing interest in digital libraries that support search for information by geographic location, or *geolibraries* (NRC, 1999).

Finally, DE is not static. In an immersive digital environment, and given sufficient general knowledge in the form of models of processes, it is possible to imagine the user of DE being able to simulate future environments, by executing models of urban growth, or species extinction, or tectonic uplift, and observing their consequences for any part of the Earth. It is also possible to imagine modeling past environments, by running simulations backwards in time. In this regard, DE is seen as having immense power for education. A static DE would be a good basis for learning the facts of the Earth's special geography. But a dynamic DE would allow students to explore the processes of general geography, and their implications, in compellingly realistic form and using boundary conditions representing environments that are familiar to them. For example, a student would be able to learn about tectonic processes by modeling the appearance of California 1, 10, or 100 million years from now.

The idea of DE as a digital library of simulation models raises interesting and challenging questions. Although libraries have evolved highly effective systems for

abstracting and indexing books and journals, and although numerous organizations now offer access to substantial resources of geospatial data (e.g., the National Geospatial Data Clearinghouse, [www.fgdc.gov](http://www.fgdc.gov)), the knowledge contained in simulation models, which is arguably among our most sophisticated knowledge of how the Earth system works, remains largely uncataloged and inaccessible. In contrast to knowledge in books and journals, no infrastructure currently provides the tools to support a scientist wishing to search for, evaluate, download, and run simulation models. An infrastructure for organizing, describing, and sharing simulation models would provide a very valuable service to the Earth science community.

The speech has spawned a substantial level of interest in DE, as a search of the WWW will reveal. DE projects are under way to model tectonic and other geophysical processes, to explore Earth imagery, to deliver the services of a map and imagery library, and to teach about the Earth and its human and physical phenomena ([www.alexandria.ucsb.edu/adept](http://www.alexandria.ucsb.edu/adept)). An interagency working group has been meeting for the past two years under the auspices of the U.S. National Aeronautic and Space Administration, and a number of prototypes of DE are under development in the private sector. The first International Symposium on Digital Earth was held in Beijing in December 1999, and another is planned in Canada in 2001.

There is every indication that DE's vision of an integrated source of knowledge about the Earth will attract the attention of researchers, developers, agencies, and corporations. This paper is about the importance of discrete global grids to that vision, and is intended to introduce global grid researchers to the potential of DE. The next section discusses the role of grids in DE, in three contexts: georeferencing, indexing, and discretization. This is followed by sections on functionality, and the assessment of grid options. The paper ends with a brief concluding section.

## **DE applications of global grids**

### ***Georeferencing***

In DE all information is referenced to positions on the Earth's surface, through point locations, *footprints*, or similar mechanisms. Services must be provided to allow users to locate places, to compare different sources of information about the same place, and to define the locations associated with information being input to the system.

Many methods exist for defining location on the Earth's surface. To be useful they must define location uniquely, at least within a specified domain (e.g., the placename Springfield is unique in the U.S. only within the domain of a state); and have meaning that is widely shared within a defined information community. Latitude and longitude referencing satisfies these requirements, but is accessible only to a comparatively small information community, as is referencing by widely adopted coordinate systems such as UTM. Placenames are the dominant method in the broader community, but are subject to a number of caveats:

- domains less than the Earth are often needed to establish uniqueness; or placenames must be coupled with one or more domain names (e.g. Springfield, Illinois, USA);
- the spatial resolutions associated with placenames are often undefined or unknown;

- the meaning of placenames can be context-specific (Los Angeles may have different meanings to speakers located in Southern California, New York, and Moscow);
- alphabets, diacritical marks, and spelling variations cause problems; and
- the meaning of placenames may change through time.

Discrete global grids offer two major advantages for georeferencing. Besides being unique and domain-independent, appropriately indexed or *linearized* grids express three-dimensional location in a single string, and make resolution explicit in the length of the string (Goodchild and Yang, 1992). Grids with these properties have been adopted in many contexts, including various systems for indexing national topographic map series; regional georeferencing schemes such as Go2; and the quadtree indexes used in spatial databases (Samet, 1989).

A user of DE may not be concerned with the precise details of internal georeferencing, requiring only services that recognize various forms of georeference, deliver predictable responses based on them, and make them interoperable. Internally, however, DE should employ a system with the union of all desirable properties. Discrete global grids appear to have the edge in this respect. Their major disadvantages—lack of a large information community familiar with them, and lack of an intuitive relationship between pairs of codes and proximity—are not important for internal purposes, but would be problematic if they were exposed to most users.

### ***Spatial indexing***

An index can be defined as a mapping of information into a linear structure, such that information on a given topic always appears at the same place in the structure. Once an index is built, it can make retrieval of information on a given topic much faster, since no search has to occur because the location of information is computable. A spatial index linearizes information based on location. Besides speeding search, a spatial index provides additional advantages in spatial databases, because of certain characteristics of their applications. The probability of a request involving two locations  $x$  and  $y$  usually decreases with the distance between them, and the probability that a request for data on  $y$  will follow a request for data on  $x$  is similarly inversely dependent on distance. It was this issue that led Morton to implement an early form of quadtree indexing in the Canada Geographic Information System (Foresman, 1997). Quadtrees and R-trees are two popular forms of spatial index (van Oosterom, 1993).

Linearized discrete global grids are clearly applicable in DE, and offer substantial advantages in data retrieval and processing. Further advantages in search over distributed archives are discussed in a subsequent section.

### ***Discretization***

If one accepts the proposition that the geographical world is inherently continuous, with an infinite number of possible locations, then digital representation must necessarily abstract, generalize, or approximate. Most forms of representation involve *discretization*, or the reduction of continuous dimensions to discrete form, most commonly with respect to the spatial dimensions. Tesselations thus replace the continuous spatial dimensions with a finite number of discrete elements, and reduce variation within elements to some

simple form, often a constant but also possibly a polynomial function, that can be described with a small number of parameters. In the modeling world the two common options are *finite element* methods, which replace continuous spatial variation with variation over a tessellation of triangles or quadrilaterals modeled using polynomials in the continuous spatial dimensions, and *finite difference* methods, which approximate spatial variation using constant values within square cells. In the geospatial data world finite difference representations are commonly known as rasters; finite element representations are known in the case of linear variation over triangular elements as *triangulated irregular networks*, and other options are generally not used. Not all geospatial data are discretized in the spatial domain—much representation of global climate data for modeling purposes relies instead on discretization in the spectral domain.

If DE is to implement simulation models, consideration must be given to the discretization inherent in its representations. Finite difference methods require regular tessellations, and the square tessellation that is normally used can only exist on a flat surface. The basis for estimating surface derivatives in GIS similarly depends on a raster representation and a flat projection of the Earth's surface. For example, the equations shown below give the standard estimation method for the first derivatives using finite differences. Discrete global grids such as QTM (Dutton, 1998), based on near-regular triangular tessellations, are more compatible with finite element methods.

$$f'_x \approx [-z_{i-1,j-1} - 2z_{i-1,j} - z_{i-1,j+1} + z_{i+1,j-1} + 2z_{i+1,j} + z_{i+1,j+1}]/8d$$

$$f'_y \approx [-z_{i-1,j-1} - 2z_{i,j-1} - z_{i+1,j-1} + z_{i-1,j+1} + 2z_{i,j+1} + z_{i+1,j+1}]/8d$$

Finally, there are obvious advantages to adopting a single approach to discretization within DE. Hierarchical methods that allow the same structure to be used over a wide range of spatial resolutions, such as the quadtree or QTM, would provide a single, consistent approach. But if QTM were to be adopted for DE, simulation models that exploit finite differences would have to be rewritten to use the finite element approach. Alternatively, a hybrid approach could be devised that would estimate derivatives for the irregularly shaped elements of QTM, but use the finite difference approach of assuming constant values within elements. For example, the derivatives at the center of a triangular element could be estimated using the approach on which the equations above are based, but adapted to the irregular geometry of QTM, by fitting a plane through the values in the element and its three full neighbors (analogous to estimating derivatives using a cell and its four full neighbors in a raster), or through the values in the element and its full and diagonal neighbors (a variable number in the QTM scheme, but always equal to 8 in a raster).

## Functionality

Several aspects of DE functionality bear directly on the topic of this conference. Reference has already been made to questions of indexing and discretization, which underlie functionality. The choice of underlying structure is also likely to be dictated by a number of other aspects of functionality.

First, DE must support a full range of georeferencing schemes, geodetic datums, and projections. Data for DE will be drawn from a wide range of sources, and merging data at global scales always raises geodetic questions, particularly when non-global datums are involved. As noted earlier, there are obvious advantages in a structure that offers consistency across a wide range of scales; a single address for any global location; and positional accuracy and spatial resolution that are explicit in the coding scheme.

Second, DE must support rapid zoom, and the Gore speech talks about zooming in real time from global to neighborhood resolutions. On a commonly available display unit of approximately 1000 by 1000 picture elements a full display of the Earth requires data with spatial resolution in the 10km range. Neighborhood display is likely to require 1m spatial resolution or better, implying that DE must support zoom over at least 4 orders of magnitude. Moreover, a complete coverage of the Earth's surface at 1m resolution requires approximately  $10^{15}$  data elements. Hierarchical structures such as QTM have inherent advantages here, since they provide a single representational structure for unlimited ranges of resolution, and adapt easily to incomplete coverage, or coverage with variable resolution.

Third, DE must support visualization. Much high-performance visualization hardware uses triangular elements, providing fast capabilities for pan, zoom, and rendering. DE displays will always be user-centered, implying that resolution can be allowed to degrade away from the center of the field of view. In this sense DE is in sharp contrast to practice in traditional cartography.

### ***Search for data in DE***

DE was presented earlier as a new metaphor for the organization of geospatial information resources. In this context, DE is envisioned as an environment in which the availability of information is discovered by the user through interaction with a virtual Earth. Information might be rendered directly onto the planet, or its existence might be represented symbolically. Prior to display, however, DE would somehow have to establish the existence of information, and determine its properties.

Currently, two mechanisms exist for discovery of such information within the distributed archive of the Internet. First, solutions such as the National Geospatial Data Clearinghouse (NGDC; [www.fgdc.gov](http://www.fgdc.gov)) rely on standards for data set description. The FGDC metadata standard (Content Standards for Digital Geospatial Metadata) provides such a standard. Participants in the clearinghouse provide metadata records for all of the data sets that are available on their servers, and users are able to search these records through a common portal. Geographic footprints are commonly expressed as bounding latitude and longitude values.

Second, a well-informed user might be able to narrow the search to a small number of servers. Many users of geospatial data know where to look for standard types of data, such as Landsat imagery, or digital raster graphics. This knowledge might be described as *collection-level metadata* (CLM), providing the general characteristics of the data sets stored on a given server or collection, in contrast to *object-level metadata* (OLM), which is used to describe the characteristics of single data sets. At this time no standards for CLM exist. In practice, many collections are defined by theme, or by data type. The Alexandria Digital Library's CLM is particularly complex, since the collection is based

on the imagery, maps, and data sets in the collection of the University of California, Santa Barbara's Map and Imagery Laboratory, one of the world's largest.

One possible solution to the CLM problem is to base server contents on a geographic tessellation. If we could agree, for example, that all data sets whose footprints intersect Santa Barbara County should be available on a server maintained by the county, then search for such data would be vastly easier than it is under the current arrangements, with such data scattered over a large number of servers maintained by different agencies, institutions, and jurisdictions. Thus discrete global grids provide a potential solution to the CLM problem, by forming the basis on which data could be allocated to a network of servers.

## **Conclusion**

At this point DE is little more than a vision. As such, however, it may be extremely useful in providing researchers and developers with a target for activities. Current technology is not sufficiently powerful to deliver all DE services, but we can certainly implement some services using current technology, and imagine a time when more powerful technology would be available. Moreover, the Vice President's vision includes significant references to central facilities—in good geographic tradition, if a service is too expensive for individuals to afford, yet still of substantial value, then it makes sense for the community to provide it as a public service at a central location.

DE can clearly benefit from developments in discrete global grids, which can provide the georeferencing, the indexing, and the discretization needed for geospatial data sets. They have properties, in particular hierarchical structure, uniqueness, explicit representation of spatial resolution, and consistency, that make them superior to any single alternative. Moreover, discrete global grids based on triangles have distinct advantages for high-performance visualization and for finite-element modeling. Of course DE must be capable of working with alternatives, for input and output, or for narrowly defined applications within the general DE structure, such as modeling on projected flat surfaces using finite differences.

Discrete global grids have many overlapping motivations, from uniform sampling of the Earth's curved surface to optimal locations for facilities. This paper has explored one more: the vision of a unified repository of all that is known about the Earth's surface and near-surface. The vision of Digital Earth appears to be capable of attracting significant interest, in communities that range from vendors of high-end visualization systems to the military.

## **Acknowledgment**

The Alexandria Digital Library and the Alexandria Digital Earth Prototype (ADEPT) are supported by the National Science Foundation and by other agencies participating in the Second Digital Library Initiative.

## References

- Dutton, G. (1998) *A Hierarchical Coordinate System for Geoprocessing and Cartography*. Lecture Notes in Earth Science 79. Berlin: Springer-Verlag.
- Foresman, T.W., editor (1997) *The History of Geographic Information Systems: Perspectives from the Pioneers*. Upper Saddle River, NJ: Prentice Hall PTR.
- Goodchild, M.F., and S. Yang (1992) A hierarchical spatial data structure for global geographic information systems. *Computer Vision, Graphics and Image Processing: Graphical Models and Image Processing* 54(1): 31–44.
- National Research Council (1999) *Distributed Geolibraries: Spatial Information Resources*. Washington, DC: National Academy Press. Also available at [www.nap.edu](http://www.nap.edu).
- Oosterom, P.J.M. van (1993) *Reactive Data Structures for Geographic Information*. New York: Oxford University Press.
- Samet, H. (1990) *The Design and Analysis of Spatial Data Structures*. Reading MA: Addison-Wesley.