

## **Historical Gazetteers: An Experiential Perspective<sup>1</sup>**

**Ruth Mostern**

**School of Social Sciences, Humanities and Arts, University of California, Merced  
rmostern@ucmerced.edu**

*For Presentation at the American Historical Association, January 5, 2006*

**DRAFT: NOT FOR CITATION WITHOUT PERMISSION OF THE AUTHOR**

### **1. Introduction: Scholarly Primitives and Spatial History**

In an influential 2000 presentation, humanities computing pioneer John Unsworth introduced the proposition that the software tools developed for the humanities should incorporate unique functionality reflecting *scholarly primitives*, or “basic functions common to scholarly activity across disciplines, over time, and independent of theoretical orientation.” His list of primitives—professed simply as a starting point for discussion—included Discovering, Annotating, Comparing, Referring, Sampling, Illustrating and Representing.<sup>2</sup> Unsworth’s paper does not refer to geographical information, and I am not convinced that his list of primitives adequately characterizes practices in historical geography (or spatial history), with or without GIS. I open my paper with Unsworth’s observation merely as a point of departure for a call to the historical GIS and digital cultural atlas community to pay similar attention to the relationship between the

---

<sup>1</sup> Acknowledgments: This paper is based substantially on research conducted at the Electronic Cultural Atlas Initiative (ECAI) at UC Berkeley between 2001 and 2004. I would like to thank the following individuals for significant contributions to my thinking on this topic: Lex Berman, Ben Brose, Michael Buckland, Linda Hill, Karen Kemp, Lewis Lancaster, Susan Stone, Zifu Zhao. Much of the research for this article was funded by two grants: “A Multilingual Gazetteer System for Integrating Spatial and Cultural Resources” (NSF IM/ITR 0114019, 2001-2002, PI Lewis Lancaster) and “Going Places in the Catalog: Improved Geographical Access” (IMLS National Leadership Grant for Libraries, LG-02-02-0035-02, 2002-2004, PI Michael Buckland).

<sup>2</sup> John Unsworth, “Scholarly Primitives: What Methods do Humanities Researchers Have in Common, and How Might our Tools Reflect This?” part of a symposium on “Humanities Computing: Formal Methods, Experimental Practice” sponsored by King’s College, London, May 13, 2000. Accessed: <<http://jefferson.village.virginia.edu/~jmu2m/King.5-00/primitives.html>>, November 8, 2005.

disciplinary traditions of history and historical geography and the creation of historical geographical information systems.

This paper concerns digital gazetteers rather than geographical information systems *per se*. As I will explain more fully below, gazetteers—indexes of named places—are, implicitly or explicitly, a foundational element of historical GIS design. Much geographically focused historical research, particularly on eras prior to the nineteenth century, is based on sources that are saturated with place names, but ones that can be assigned geographical coordinates only with difficulty, and with sometimes undesirable loss of the ambiguity inherent in the original text. Think for example, about narratives of war or travel. The writing of history in narrative form allows us to skirt these issues, but when we create geographical information systems, we are forced to take a stand on the locations of places and the events that occurred in them. Gazetteer databases are well equipped to incorporate notes, certainty ratings, alternative names in multiple languages, feature types, relationships among places, and citations to sources in ways that are reasonably familiar to historians. Furthermore, gazetteers are the only plausible starting point for historical geographical research deriving from texts, for which the first step is to identify named places and then to record some information about them. It is my hope that articulating design practices for gazetteers that incorporate thick information about named places in the past will also reveal and reflect scholarly primitives for historical geography.

## **2. Digital Gazetteers**

In a classic 1975 article from which the title for this paper is taken, cultural geographer Yi-Fu Tuan explicitly denies the capacity of gazetteers to encompass information about places from a humanist perspective. He writes: “Geographers have approached the study of place from two main perspectives: place as a location, a unit within a hierarchy of units in space; and place as a unique artifact...[In the second sense,] place is a center of meaning constructed by experience.” He continues: “Experience constructs places at different scales. The fireplace and the home are both places. Neighborhood, town, and city are places; a distinctive region is a place, and so is a nation...They are all centers of meaning to individuals and to groups. As centers of meaning, the number of places in the world is enormous, and cannot be contained in the largest gazetteer.”<sup>3</sup> I am introducing this quote as a bit of a red herring. After all, Tuan mentions gazetteers in this article only in passing, and is presumably concerned with an older meaning—the list of place names in the back of an atlas—and not the relational databases, XML markup, networked services, and map interfaces that I have in mind here. Nevertheless, he elegantly raises precisely the question that I wish to address. Namely, is there such a thing as a humanistic historical gazetteer?

In a widely accepted definition developed at the pioneering Alexandria Digital Library (ADL), a digital gazetteer includes three required elements. It is a list of proper names of geographic places and features, together with at least one geographical footprint (in the form of a coordinate, bounding box, or completely polygonal boundary), and at least one feature type (city, volcano, temple, empire) associated with each named place. In this sense, the ADL developers note, a gazetteer is an important reference work in its own right. It answers “where is” questions, translates between geographic names and

---

<sup>3</sup> Yi-fu Tuan, “Place: An Experiential Perspective,” *Geographical Review* 65.2 (April 1975).

locations, and allows users to locate particular types of geographic features in a designated area.<sup>4</sup> For humanists, some of the results that can be readily acquired even with existing content are of notable research interest. For instance, a search of the Alexandria Digital Library for the name of my hometown Merced brings up 80 matches throughout California, Latin America and the Philippines, instantaneously producing a rough map of the Spanish cultural world. A historical source like the gazetteer-based China Historical GIS, once complete, will be able to respond to a query about the city of Beijing 北京 by mapping the locations of the approximately two-dozen places that have at one time been designated as a northern capital—the translation of Beijing—in the Sinophone world.<sup>5</sup>

Alexandria Digital Library Gazetteer Server Client. Using ADL Server Use ESRI Server

Search Results: 80 matches (Problem Report)

#	names	Class
1	La Merced - Cundinamarca, Departamento de - Colombia	populated places
2	La Merced - La Rioja, Provincia de - Argentina	populated places
3	Fort Funston (historical) - San Francisco County - California - United States	historical sites
4	La Merced - Huancavelica, Departamento de - Peru	populated places
5	Merced Inuez - Camaguey, Provincia de - Cuba	populated places
6	Nuestra Señora de la Merced - Ilocos Norte, Province of - Philippines	populated places
7	La Merced - Canar, Provincia del - Ecuador	populated places
8	Quebrada La Merced - Pichincha, Provincia de - Ecuador	populated places
9	La Merced - Guanajuato, Estado de - Mexico	populated places
10	Merced Milagro - La Rioja, Provincia de - Argentina	populated places
11	Merced Gomez - Federal, Distrito - Mexico	populated places
12	Merced - Ecuador	populated places
13	Merced - Granma, Provincia - Cuba	populated places
14	Hacienda La Merced - Imbabura, Provincia de - Ecuador	populated places
15	Hacienda La Merced - Chimborazo, Provincia del - Ecuador	populated places
16	La Merced - Veracruz-Llave, Estado de - Mexico	populated places
17	La Merced de Buenos Aires - Imbabura, Provincia de - Ecuador	populated places
18	Hacienda La Merced - Pichincha, Provincia de - Ecuador	populated places
19	La Merced - Canar, Provincia del - Ecuador	populated places
20	La Merced - Ancash, Departamento de - Peru	populated places
21	La Merced - Coahuila de Zaragoza, Estado de - Mexico	populated places
22	Hacienda La Merced - Pichincha, Provincia de - Ecuador	populated places
23	La Merced - Veracruz-Llave, Estado de - Mexico	populated places
24	La Merced - Santa Fe, Provincia de - Argentina	populated places
25	La Merced - Olancho, Departamento de - Honduras	populated places
26	La Merced - Tarija, Departamento de - Bolivia	populated places
27	Hacienda La Merced - Pichincha, Provincia de - Ecuador	populated places
28	La Merced - Catamarca, Provincia de - Argentina	populated places
29	Merced (historical) - Poinsett County - Arkansas - United States	populated places
30	Hacienda Merced - Azuay, Provincia del - Ecuador	populated places
31	Hacienda La Merced - Tungurahua, Provincia del - Ecuador	populated places
32	La Merced - Guayas, Provincia del - Ecuador	populated places
33	Merced de Capayan y Alieca - La Rioja, Provincia de - Argentina	populated places
34	Merced de Tudum - La Rioja, Provincia de - Argentina	populated places
35	La Merced - Chihuahua, Estado de - Mexico	populated places
36	Merced - Santiago de Cuba, Provincia de - Cuba	populated places
37	Merced Falls - Merced County - California - United States	populated places
38	Hacienda La Merced - Pichincha, Provincia de - Ecuador	populated places
39	Merced La Pampa - La Rioja, Provincia de - Argentina	populated places
40	Hacienda La Merced - Chimborazo, Provincia del - Ecuador	populated places

Please set at least one search condition.

Location:  within map  overlaps map  anywhere

Place Name:

Feature Type: ( Feature Type Thesaurus)

-Any Type-  
General Categories  
hydrographic features  
land parcels  
manmade features  
physiographic features  
regions  
administrative areas  
military areas

Place Status:  ANY  Current  Former  Proposed

Identification Code:

Comments to: [ADL Gazetteer Development](#)

This service hosted by  
Alexandria Digital Library, Davidson Library, Univ of  
California, Santa Barbara

Gazetteer Client Copyright © 2003-2004. The Regents of the University of California. All Rights Reserved.  
Uses the Gazetteer Protocol. Server and client code can be downloaded from:  
[Alexandria Digital Library Gazetteer Development](#)  
Mapping powered by ESRI ArcIMS technology

<sup>4</sup> Linda L. Hill, James Frew and Qi Zheng, “Geographic Names: The Implementation of a Gazetteer in a Georeferenced Digital Library,” *D-Lib Magazine* 5.1 (January 1999) <<http://www.dlib.org/dlib/january99/hill/01hill.html>>, accessed December 10, 2005. Also cite to the ADL content standard and say something about non-required and repeatable elements—this should probably be in the text, anyway, not the footnote.

<sup>5</sup> China Historical GIS, see <<http://fas.harvard.edu/~chgis>>, accessed December 18, 2005.

**Figure 1: The Alexandria Digital Library interface, showing the results of a search for the place name Merced.**

Gazetteers are exceptional tools for place name reference and simple queries. However, that is not their only function. They also provide the basis for place-based search, display and integration. Any type of information that includes the name of a place can be associated with a gazetteer: texts, images and remotely sensed data; library records or databases; dates of events associated with places; alternative names in multiple languages (which may each have their own starting and ending dates), and so on. In this capacity, gazetteers can serve as super-powerful core element in any information architecture. They can form the basis for a spatial visualizations, including even animated historical maps, or more broadly, for digital atlases that incorporate demographic information or links to museum objects along with maps. They can permit map-based library searches that return references to any resources pertaining to places within a designated bounding box, even if the searcher does not know the names of the places she is interested in, only the region on the map.<sup>6</sup> Commercial services like the ubiquitous GoogleMap are based on gazetteer architecture.

Finally, network service protocols allow multiple gazetteers to be searched and used together. As Unsworth contends in the “Scholarly Primitives” article, “the most interesting things that you can do with standalone tools and standalone resources is, I would argue, less interesting and less important than the least interesting thing you can do with networked tools and networked resources.”<sup>7</sup> Historical gazetteers are particularly likely to be specialist resources, and developing a networked environment for them will

---

<sup>6</sup> For ongoing research on georeferencing for library catalogues, see Michael Buckland’s “Support for the Learner: What, Where, When and Who,” <<http://www.ecai.org/imls2004/index.html>>, accessed December 18, 2005.

<sup>7</sup> Unsworth, “Scholarly Primitives.”

be of particular importance. This point is primarily the topic for another article. In this context, I simply note that the desirability of networking multiple historical gazetteers underscores the importance of developing shared standards.

### **3. Digital Gazetteers and Scholarly Primitives for History: Research and Challenges**

Gazetteers have the capacity to meet Yi-fu Tuan's call for a data source that represents places as structures of meaning—but only in theory. I return to the question of the current capital of China as an example of some issues that historical gazetteers must confront. In the search that I proposed for historical references to northern capitals of China, I called the place that I was interested in “Beijing,” which presumed Pinyin transliteration. A gazetteer would also have to include the term “Peking,” the same place rendered in the Postal transliteration, and “Pei-ching,” the same place in Wade-Giles, as well as the Chinese characters, with care to render them in UNICODE to facilitate cross-platform display. If I were Mongol I would have known the current capital of China—one of many Beijings to have existed historically--as Khanbaliq in Mongol or as Dadu 大都 in Chinese; if I were the medieval European traveler John Mandeville I would have called it Gaydon. A die-hard anti-Communist in Taiwan will still call it Beiping 北平 today, exchanging the term for “capital” (*jing*) with the one for “peace” (*ping*) to deny the legitimacy of the current regime. A medieval poet, writing about the same city before it served as an administrative seat, would have known it by its classical name of Yan 延. All this for the name of a single place: surely this is experiential geography! It is also information that can be structured in a digital gazetteer, but doing so requires several

developments if it is to reflect historians' primitives. The remainder of this paper focuses specifically on good practices for historians modeling historical spatial information.

A historical digital gazetteer *for historians* must be a useful digital resource, taking advantage of the capacities offered by a computing environment for structured query and analysis, resource integration, and cartographic visualization. However, it must also reflect the lived experience of space by people in the past, and the primitives familiar to contemporary historians who document that experience. It is in this attempt to address the requirements and cultures of both history and information management that humanities computing finds its most intriguing and robust expression. In the remainder of this paper, I introduce some of the requirements for historical digital gazetteers that conform to both cultures, with examples from my own research on early modern China.

1. The capacity to model complex information about places. As in the Beijing example above, a single place may have multiple names. Each name—not just the place itself, but each moniker for it—has a linguistic affiliation, a start- and end- date, and a political affiliation. But historical spatial information can become far more complex than that. To cite an example from my own research: when the two provinces of Guangnan 廣南, East and West, came under control of the Song dynasty after 971 CE, over 200 counties and prefectures (the latter were units larger than counties and smaller than provinces) were abolished within the space of a few years. The process by which a new spatial organization in southeast China came into effect was complex.

- In some cases, counties and prefectures were abolished, re-established and re-abolished within the space of months.

- In others, changes were only proposed but never really enacted, and it is often difficult to determine at 1,000 years distance which were which.
- Sometimes, as places were consolidated, their names were changed as well.
- Once abolished, some counties and prefectures saw the entirety of their territory engulfed by a neighboring jurisdiction, while others were split between several units.
- In the case of prefectures, some were abolished altogether, while others were merely demoted to counties.
- In some cases, the county seat was moved from one town to another, and in others the total footprint of an existing territory changed even if its name and hierarchical status remained the same.
- After the mass of abolitions was completed, the remaining counties were traded among a new and reduced landscape of prefectures.<sup>8</sup>

A database cannot fully replace a monograph that explains the intricate and surprising political forces that motivated these vast changes and the regime changes that motivated them. However, it can model them, it can be used for research and query to test hypotheses about their spatial and temporal distribution, and it can provide the content for exceptional maps to illustrate them. But if it is to do so, it needs to structure temporal change in great detail. As this example illustrates, temporal information about geographical objects in the past is not just a yes-no answer about the existence of the place as a whole. In order to accurately model the spatial history of tenth century Guangnan, dates must be associated with each name of each place, each feature type

---

<sup>8</sup> Ruth Mostern, "Dividing the Realm in Order to Govern: Spatial Organization and State Power in the Tang-Song Transition," forthcoming in Peter Lorge, ed., *Between and Among Empires: Ideological Space in the Tang-Song Transition* (Hong Kong University Press, forthcoming 2006).

(county or prefecture), each location (as represented by its boundaries, constituent units and county seat), and its parent-child relationships with other jurisdictions.

2. The capacity to incorporate rich documentation. Some of the items in literary critic John Unsworth's list of primitives are of interest primarily to scholars concerned with textual analysis. The one that resonates most with the practice of historians is what he calls *Referring*. Historians writing about spatiality, or any other subject, for a university press, a dissertation committee, or a scholarly journal, are accustomed to producing massive infrastructures of citation and documentation. As is typical for historical writing, my eight page discussion about tenth century Guangnan in a recently completed paper includes fifteen footnotes to tenth century sources documenting the territorial changes that I discuss, to other contemporary historians' interpretations of the era in question, and to explanations of how I reached particular conclusions. Digital gazetteers that are to be used with confidence as sources for historical research or cartographic display by historians need to incorporate the equivalent of footnotes. A historical gazetteer must be considered a richly bibliographical tool. The model for each element in a gazetteer entry about a given place--the entry for a given location as a whole, each alternative place name, each feature type, each spatial location, and the starting and ending dates of each of these pieces of information—must allow for attaching references to primary sources, scholarship, digital images, notes, or other information that allows working historians to approach the gazetteer database with the same confidence that they would bring to a peer-reviewed monograph on the same topic.

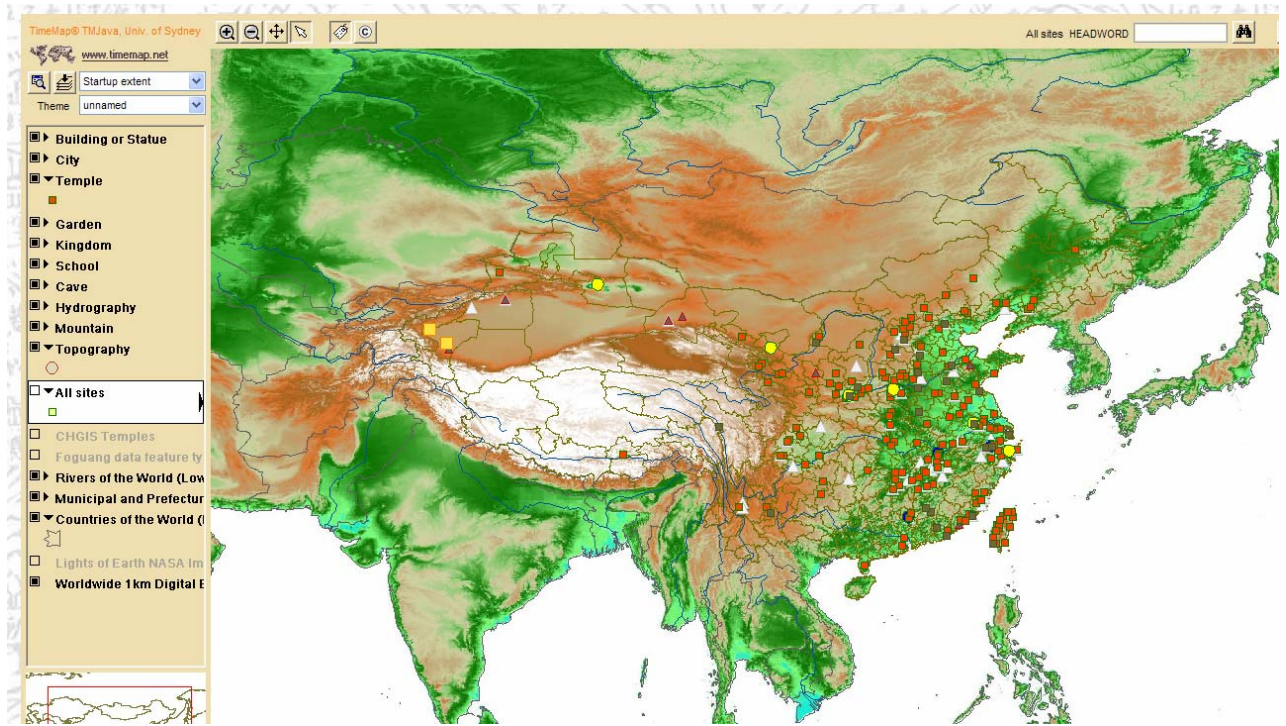
3. The capacity to document uncertainty. Research on uncertainty in GIS is an intriguing and highly technical field that involves collaboration between geographers, professional mathematicians, computer scientists and practitioners of fuzzy logic.<sup>9</sup> Following Unsworth's argument that effective digital tools for humanists are ones that reflect indigenous scholarly practices, it is my contention that historians creating digital gazetteers need not concern themselves overmuch with those findings. Surely, spatial historians have home-grown traditions for managing uncertainty! After all, the books and articles we write are devoted in large part to high-level discussions about the contingent nature of our interpretations, ranging from evaluations of the ideological predilections of our sources, to analyses of competing territorial claims by different regimes, to notes about a complete lack of documentation for some topic of interest.

Uncertainty, ambiguity, silence, and contradiction are common to almost all areas of historical research. For spatial history in specific, geographical and temporal information are particular challenges. Let me provide one more example from my own research. In 2003-4, I was the lead researcher for the Digital Atlas of China and the Himalayas, a Luce Foundation funded project under the direction of Lewis Lancaster. As a first step, the project team extracted the names of approximately 1,000 Buddhist-associated places in China from the *Foguang Encyclopedia* (*Foguang dacidian* 佛光大辭典), a major digital reference work for Buddhist studies.<sup>10</sup>

---

<sup>9</sup> A frequently cited book on this topic is Peter Burroughs and Andrew Frank, eds., *Geographic Objects with Indeterminate Boundaries* (London: Taylor and Francis, 1996).

<sup>10</sup>*Foguang dacidian*. Available on-line at <<http://www.hsilai.org/Chinese/SearchDict.asp>>, accessed December 18, 2005. A draft website for the Digital Atlas of China and the Himalayas is available at <<http://ecai.org/chinareligion>>, accessed December 18, 2005.



**Figure 2: Locations of Buddhist-related sites in China from the *Foguang Encyclopedia*. Note the distribution into various feature types.**

A typical entry in the *Foguang Encyclopedia*, for Mount Emei 峨眉山 in Sichuan province, one of four sacred mountains in Chinese Buddhist geography, references 25 by name in addition to the mountain itself, and another 210 places by category. The entry includes the following information:

- The location of Mount Emei in political space (province and county).
- Four alternative names (including Buddhist and Daoist names).
- Neighboring topographical features.
- Five of its named peaks.
- Its status in the corporate group of the Four Great Sacred Mountains, all of which are named.
- Significant named built features on the mountain: five temples (one with both a historical and contemporary name), one statue, one road.

- Other referenced features include more than 70 additional temples, more than 40 grottoes and caves, and more than 100 stone niches for carvings.

The *Foguang Encyclopedia* entry for one of the famous temples on Mount Emei includes the name of an additional corporate group (The Six Great Ancient Monasteries of Mount Emei), three historical names for the temple, seven named buildings, a named statue, and a named relic.<sup>11</sup>

An “experiential” historical gazetteer of Chinese religion would surely need to include entries for all of these entities. For the purpose of this discussion of uncertainty, several points are noteworthy. First, none of these entities are georeferenced. Temples, for instance, are simply described as being located on the mountain. A significant source of ambiguity in a historical gazetteer arises simply from the use of textual sources that do not have a tradition of precise geographical reference. For the initial draft of the Digital Atlas of China and the Himalayas, we have simply referenced all of the places on Mount Emei to the mountain itself, with the hope of creating more precise information based on additional written sources or ground-truthing as funding permits. Only then will we have created an experiential gazetteer that represents the ways that pilgrims, monks, tourists, donors, artisans and others have occupied Mount Emei. In the meantime, the draft gazetteer needs to be structured such that the paradigms we have used to record approximate locations of entities on the mountain can be adequately documented.

For most historical gazetteer projects, temporal information is even more of a challenge than spatial information. The requirements of a digital gazetteer that can drive map visualization are entirely different from those of an encyclopedia, let alone the

---

<sup>11</sup> *Foguang Encyclopedia*, <<http://www.hsilai.org/Chinese/dictsub.asp?DINDEX=12680>>, accessed December 18, 2005.

primary sources that historians use or the monographs that we write. As a case in point, the entries for the temples, grottoes and other human-built places on Mount Emei that are listed in the *Foguang Encyclopedia* include only names with no dates of founding. In many cases, only names are listed, with no founding dates. When information about founding is recorded, it often references dynasties that lasted for centuries, not particular years. In almost no case does the *Encyclopedia* describe, even approximately, the date when a site fell into disuse. It does, however, sometimes list both a date of founding and a date of rebuilding, thus making it clear that the site was not continuously occupied.

To provide fidelity to historians' primitives, to create gazetteers that reflect lived experiences in the past, and to facilitate meaningful map display, these ambiguities need to be rectified. I do not believe that this kind of ambiguity can or should be modeled algorithmically as many in the more mathematically inclined "fuzzy GIS" community advocate. Careful data modeling is the solution that best fits both the characteristics of historical spatial data and the community practices of historians. If unknown starting and ending dates are represented by *EarliestPossible*, *LatestPossible* and *BestGuess* dates, and each of those supports bibliographical fields and notes fields, then it is possible to create a usable resource while doing digital history well. Users would have the option of selecting relatively more liberal or more conservative date ranges based on their needs.

#### **4. Conclusion: A Historian's Proposal for Gazetteer Design Primitives**

A number of scholars have been addressing issues of how to develop gazetteers that better reflect the disciplinary practices of historians and humanists. The same special issue of *D-Lib Magazine*, edited by Alexandria Digital Library researcher Linda Hill,

which reports the results of the gazetteer projects carried out under the auspices of the Electronic Cultural Atlas Initiative, also includes an article by Gregory Crane, founder of the Perseus Project, one of the first and most influential spatially referenced digital libraries for the humanities. Crane, arguing that various cultural documents describe space in very different ways, begins his article with the observation that “While space has an objective component, cultural practices of describing space naturally differ. Ancient Greek and Latin texts, for example, do not draw upon complex geographic sources available to modern documents but their naming systems are much more well defined and these texts are much more tractable to machine processing than modern US sources.”<sup>12</sup> As Yi-fu Tuan also argues, it is this human experience of space that gazetteers and other digital information resources need to capture. As Crane notes, local systems of knowledge even now are subjugated to the needs of nation-states. He is concerned that creating systematic, geo-referenced data for a gazetteer risks making researchers less capable of appreciating the meaning of a landscape to, for instance, the foragers who first inhabited it.

As I have argued in this paper, excellent gazetteer design can reveal, rather than conceal, lived experiences of space. A foragers’ gazetteer can be developed that lists sites of meaning, power and utility in a landscape. Gazetteer developers simply need to commit to depicting the types of places, relationships among them, and frequent geographical changes that characterize meaning constructed by experience.

Gazetteers are not intended to describe places in the same ways that pilgrims, shamans or poets do. However, they are an essential component of any historical

---

<sup>12</sup> Gregory Crane, “Georeferencing in Historical Collections,” *D-Lib Magazine* 10.5 (May 2004). <<http://www.dlib.org/dlib/may04/crane/05crane.html>>. Accessed December 10, 2005.

information system, and it is therefore incumbent on us to develop design principles that are consistent with our disciplinary practices and the requirements of our research.

Developers of gazetteers must think broadly about the kinds of projects that systematic, georeferenced geographical information is good for, including experiential cultural atlases, and not just the lists of administrative jurisdictions that are most often associated with gazetteers.

In this spirit, I end in response to John Unsworth by proposing a set of “scholarly primitives” for historical geography and spatial history: **Attributing** sources extensively, **Historicizing** temporal change in spatial organization, **Contextualizing** the social, cultural and political basis for spatial organization, and **Modeling** uncertainty and ambiguity in time and space.