# Spatial Indexing for Document Access; Using a GIS to Index Documents

## Introduction

Spatial indexing is based on the physical grouping of items that are indexed. For example, countries can be grouped by continent. Telephone books group telephone numbers by city. Claudius Ptolemy organized his astronomical observations spatially and provided the basis for 15 hundred years of spatial organization for celestial materials. City records are often grouped by the parcel of property to which the records pertain. With spatial indexing, it is even possible to provide VR (virtual reality) fly throughs (or fly overs, or fly unders) to help users maintain the context of their search as the users expand or narrow the scope of their search. Spatial indexing has a richness that parallels indexing by subject or author.

Full text indexing has been one of the few new things to come from the introduction of computers. Everyone is hoping for at least one more clear benefit of the computer age. Speech recognition may be one. Matching images by content seems to be much more elusive. The idea of spatially indexing information, based on a geographic location, may be as powerful (and as tied to computers) as full text indexing has been.

In spatial indexing for records management, records are found based on their association with a place. Many records are tightly coupled to a place. Land records and most municipal government records are good examples.

Fiction could be indexed geographically based on the locating in which a story takes place.

Like other forms of indexing, geographic indexing can be combined with other indices. For example, one might as for all books of fiction that take place in upstate New York in the nineteenth century and that were written in the twentieth century.

USC has a project on the Internet in which archival records are indexed geographically. The project URL is: <a href="http://www.USC.edu/isd/locations/cst/IDA/prospectus.html">http://www.USC.edu/isd/locations/cst/IDA/prospectus.html</a>

The City of Los Angeles GIS (Geographic Information System) is a highly integrated set of spatial (geographic) indices, including a composite and digitally rectified (the pixels are in the right place) digital orthophoto [http://mapping.USGS.gov/www/ndop] to represent images of things (such as buildings) in their true location. (The City is using a specially flown and processed image rather than the USGS image.)

The website given below gives everyone a chance to try out a fully developed spatial index. This index could be easily integrated with the USC project. For example, both systems use the same <u>http://www.ESRL.com</u> (Earth Sciences and Resources Institute) software and cover approximately the same geographic area. Linking the two would be something like cataloging books by looking up the Library of Congress classification (that is, easy and deterministic, yielding the same results for almost anyone doing the process).

Unlike traditional library cataloging, spatial indexing receives funding from sources such as real estate, civil engineering, environmental planning, disaster planning, and tax collection. Spatial indexing provides a method of synergistically extending the benefits of indexing expenditures, provided by various business activities, into other areas such as the humanities.

Eventually all materials will be spatially cataloged, and the richness of our record will have a new key (spatial indexing) to unlock access and provide a new perspective

Of course, spatial indexing is not for all books (or other items). Forcing spatial indexing when there is no identifiable geographic location is similar to insisting that every book be cataloged by author, even when no author

can be identified for a given book. Because spatial indexing is now computer based, the existence of a spatial index need not impact the searches of people who have no use for spatial indexing.

## The City of Los Angeles Spatial Index

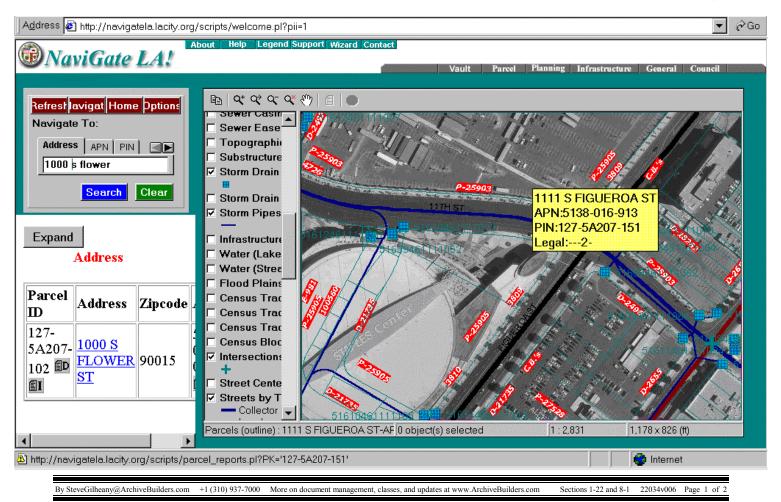
The City of Los Angeles has placed just over 700 thousand images on the Internet. The system is also designed to support as a map layer, a digital orthophotographic image (areal photograph) that will provide a continuous tone image for the entire City. Both the contone digital orthophoto and the more than 250 vector data GIS layers will be available at continuously variable map scales to index documents, and to provide a simple version of a VR (Virtual Reality) fly over of the City.

The system also supports interactive submission of large format images for the redlining of plans and the issuance of permits over the Internet. The redlining capability can also be used by field workers to input data, via the Internet, that can be used to create real time maps of field incidents, events, and status on the Internet. With this feature, users can provide system indexing. This is similar to libraries adding printed bibliographies to their collections, but the field works can do spatial indexing in real-time, in response to events such as earthquakes, mudslides, floods, and fires.

The website for presented for the GIS indexed documents and digital orthophoto is <a href="http://NavigateLA.LACity.org">http://NavigateLA.LACity.org</a>

# Navigate to Staples Center

The image below shows using three methods of indexing: a relational database, and GIS vector data overlaid on a raster based areal photograph. The P and D numbers are document index numbers linked to a location on the map.



# GIS Indexed Engineering Drawing Repository on the Internet with Markup: City of LA Case Study

# Note to Readers

#### **Updates and More Detailed Descriptions**

When using the information in this article, please check the website www.ArchiveBuilders.com for updates. The version number of this article is just before the page number below. The website also has articles that provide more details on some of the terms and concepts in this article.

#### Comments

Please let us know how you like this paper, or if you had any questions. What would you like to see in the future? For more, and the most recent version of this article, please visit our web site at www.ArchiveBuilders.com.

Please send your comments via email to SteveGilheany@ArchiveBuilders.com. Tel: +1 (310) 937-7000. Fax: (310) 937-7001. Also, please let us know where you saw this article.

### Acknowledgements

Reprinted from <u>Archive Planning</u>, Volume 5, number 8, 2001, Archive Builders' analysis newsletter for document management.

See www.ArchiveBuilders.com.

All trademarks are the property of their respective holders.

#### Note to Editors

# Paper 22034v004

We will continue to update these articles as we get comments. Please contact us for the most current version before you publish. Also, please request permission to publish the article. Permission will be given freely for most purposes.

Steve Gilheany Archive Builders 1209 Manhattan Ave., C-14 Manhattan Beach, CA 90266 Tel: +1 (310) 937-7000. Fax: (310) 937-7001 SteveGilheany@ArchiveBuilders.com

# Bio

Steve Gilheany, BA in Computer Science, MBA, MLS Specialization in Information Science, CDIA (Certified Document Imaging System Architect), AIIM Maser, and AIIM Laureate, of Information Technologies, CRM (Certified Records Manager, ARMA) has twenty years experience in document imaging and is a Sr. Systems Engineer at Archive Builders.

#### Author

Steve Gilheany is a Sr. Systems Engineer at Archive Builders. He has worked in digital document management and document imaging for twenty years.

His experience in the application of document management and document imaging in industry includes: aerospace, banking, manufacturing, refining, natural resources petroleum transportation, energy, federal, state, and local government, civil engineering, utilities, entertainment, commercial records centers, archives, non-profit development, education, and administrative, engineering, production, legal, and medical records management. At the same time, he has worked in product management for hypertext, for windows based user interface systems, for computer displays,

for engineering drawing, letter size, microform, and color scanning, and for xerographic, photographic, newspaper, engineering drawing, and color printing.

In addition, he has nine years of experience in data center operations and database and computer communications systems design, testing, and programming, software configuration management. He has an MLS Specialization in Information Science and an MBA with a concentration in Computer and Information Systems from UCLA, a California Adult Education teaching credential, and a BA in Computer Science from the University of Wisconsin at Madison. His industry certifications include: the CDIA (Certified Document Imaging System Architect) and the AIIM Master (MIT), and AIIM Laureate (LIT), of Information Technologies (from AIIM International, the Association of Information and Image Management, www.AIIM.org), and the CRM (Certified Records Manager) (from the ICRM, the Institute of Certified Records Managers, the official certifying body for ARMA International, the Association of Managers Administrators, Records and www.ARMA.org).

# Contact:

SteveGilheany@ArchiveBuilders.com Tel: +1 (310) 937-7000. Fax: (310) 937-7001

For more information, courses, and papers: http://www.ArchiveBuilders.com