Office of Statewide Health Planning and Development

Conceptual HealthCare Facilities Location Data Model

June 30, 2002

Abstract:

Business activities and decision-making undertaken by OSHPD require location-based information. This conceptual geographic data model identifies the subject matter and subsequent data structure that enables OSHPD to incorporate mapping and spatial analysis with other existing information. Mapping location-based information requires an understanding of the relationship among the licenses that authorize healthcare services, the facilities that OSHPD recognizes, and the physical location where these reside. This preliminary model generalizes those real-world relationships in order to support an array of analyses and map-based applications.

Prepared By:

VESTRA Resources, Inc.
962 Maraglia Street
Redding, CA 96002
(530) 223-2585

CONTENTS

LIST OF FIGURES

SECTION 1—INTRODUCTION

OSHPD Business Program Background

The Office of Statewide Health Planning and Development (OSHPD) provide a critical role in improving access to quality healthcare for all Californians. A department of the California Health and Human Services Agency, OSHPD is responsible for planning and supporting healthcare delivery systems to meet the current and future needs of the people of California.

OSHPD Mission Statement

The Mission of OSHPD is to be a leader in analyzing the capacities and gaps in California's healthcare infrastructure, promoting a diverse and culturally competent healthcare workforce, informing the public about the quality of healthcare in California, assuring the safety of healthcare facilities, providing financial support for non-profit healthcare facilities, and facilitating community development of sustained capacity to address local healthcare issues.

In March 2001, following his appointment as Director, Dr. David M. Carlisle announced the office's vision to start the new millennium:

OSHPD Vision

Equitable Healthcare Accessibility for California

Several business priorities establish the basis of OSHPD's direction and goals and a framework for implementing this vision. All of the following business priorities require the ability to measure quantitative information and as a result involve data collection, organization, analysis, and display:

Access To Care: Access depends in large part on having the right healthcare resources in the right place at the right time. OSHPD addresses this priority through cooperative agreements with community-based primary care facilities and physician training in under-served areas.

In addition, programs are also in place to serve economically disadvantaged students and career re-training programs in order to meet the workforce demands of California's healthcare industry.

Healthy Communities: Community health status programs assess community-specific health issues and obtain and allocate federal and state financial support.

Assessing community health status requires measurements and information that are specific to individual communities.

Quality of Care: OSHPD oversees the public's access to objective vital health. This involves analysis, monitoring, and assessment of healthcare delivery. OSHPD also oversees improved patient safety programs through pilot projects to reduce medical error, and provides report cards on hospitals, health plans, and medical groups.

Quality assessment requires measurements and information that relate patient outcome with the healthcare facilities and the communities they serve.

Healthcare Facility Support and Development: The healthcare infrastructure of California is based on community facilities meeting local needs. OSHPD addresses this priority through programs that ensure earthquake resilience among hospitals, and also enable favorable loan opportunities for healthcare facilities to obtain financing to develop or expand their services to the communities they serve.

All of these OSHPD business priorities depend on access to and analysis of *place-based information*.

Using Geographic Information to Improve Decision-making

The unique aspect of GIS technology, relative to other more traditional tabular data analysis and reporting tools, is its ability to show and summarize relationships in the data based on geographic location. This is especially important if the tabular data sets for supporting a particular business function have been summarized for dissimilar or overlapping areas.

OSHPD is implementing an Enterprise GIS (EGIS) in order to link the information they manage to location. This technology will provide OSHPD better tools and more understandable information that will improve the policy decision-making process. Improving the policy and decision-making process will help resolve many of the underlying business process-related problems at OSHPD, and reduce the negative consequences of those problems. GIS will provide more understandable information for the four components of the OSHPD policy and decision-making process (as depicted in Figure 1):

1. Assessment of Conditions
2. Allocation of Resources
3. Measurement of Outcomes
4. Communication of Results

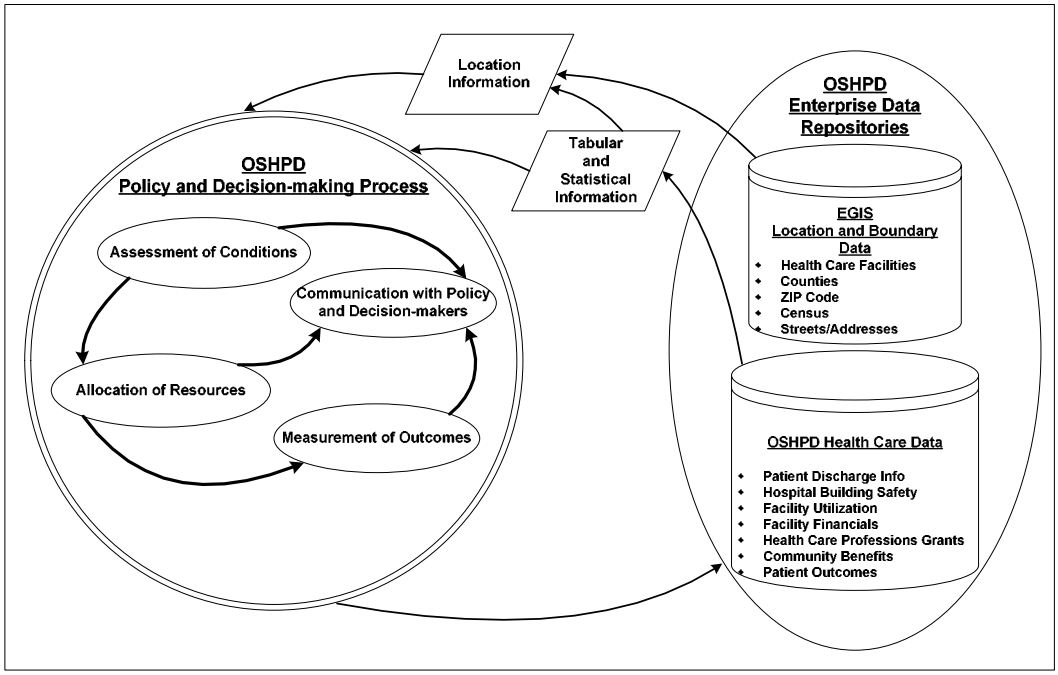

Figure 1. Using location information to improve the policy and decision-making process.

SECTION 2—CONCEPTS OF GEOGRAPHIC DATA MODELS

Overview of the Geodatabase Data Model

The primary purpose for developing a geodatabase data model is to enable features stored in GIS datasets to be "smarter". By using an object-oriented approach, GIS database subjects are modeled to have "natural" behavior, and are defined to mimic their real world counterparts. This modeling includes the general and spatial relationships that describe how objects are connected in the real world. The geodatabase data model closes the gap between physical data model and the logical model of real life features.

The geodatabase data model permits the implementation of many custom behaviors without the need to develop custom code. Examples of custom behavior are rules that limit the ranges of possible attribute values (domain ranges) and rules that test for correctness of data input (validation rules).

Benefits of using a Geodatabase

Designing data models using geodatabases provide several benefits. Those that are particularly relevant to the OSHPS EGIS and business functions are:

A uniform repository of geographic data: The framework location-based information collected and managed by OSHPD can be stored consistently in one central database.

Users work with more intuitive objects: The OSHPD Facilities geodatabase can contain objects that correspond to their own organizational model of data, with respect to specific business requirements.

Features have a richer context: One feature's context with another is defined through spatial representation and general relationships. This lets one define what happens to a feature following a change in another feature. It also permits inspection of features and their information based on the location of another related feature.

There are other benefits of geodatabases that will indirectly benefit the OSHPD EGIS by improving editing, dynamic mapping, and data access in a multiple user environment.

Extending Databases using Geodatabases

A geodatabase is not unlike a traditional relational database (RDMS). A geodatabase extends the benefits of traditional relational databases by providing additional enhancements specific to spatial data. Some of the enhancements that are relevant to OSHPD include:

1. Geodatabases can simultaneously represent spatial data as discrete objects as vector features, continuous phenomena (ie: distances) with rasters, and references to places such as addresses.
2. A geodatabase can define and store the coordinate system information for a set or sets of data.
3. A geodatabase contains the general and arbitrary relationships between objects.
4. A geodatabase can unite the natural “real world” behavior of objects to the tables that store features.

Implementing Geodatabases

There are two types of geodatabases, personal and enterprise.

Personal geodatabases are suitable for small, “project-level” implementations. They use the Microsoft Jet Database Engine and are implemented and stored as Microsoft Access databases.

For large enterprises, such as OSHPD, geodatabases are deployed using ArcSDE, a multi-user spatial data access extension to a relational database. OSHPD will be implementing an enterprise geodatabase on Microsoft SQL Server 2000.

SECTION 3—HEALTHCARE FACILITIES WITHIN OSHPD BUSINESS PROCESS

Because of the diversity of programs and associated healthcare information collected and managed by OSHPD, having a single, consistent way of representing a healthcare facility is an important, yet daunting task. Each program has its own set of criteria and characteristics that define a "facility". The one descriptive element that is common among all representations of facilities is location. Therefore, location is the single characteristic that can be used to unify a broad array of independent tabular databases that are needed to support OSHPD's business process. The following sections of this document describe the logical data "objects" and the relationships that define healthcare facilities.

Sites

There are between 6,500 and 7,000 individual locations (sites) in California that provide healthcare services. These include hospitals, long-term care facilities, primary and specialty care clinics, home health agencies, and hospices.

A ***Site*** is generally identified by a street address. The address is the primary means to differentiate one physical location from another.

A site is the physical location at which one or more healthcare service is provided (see Figure 2). For example, a site may be the address of an outpatient clinic as in "Location C", or it may be the address of a complex "campus" of services including a hospital and other buildings where numerous services are provided ("Location A").

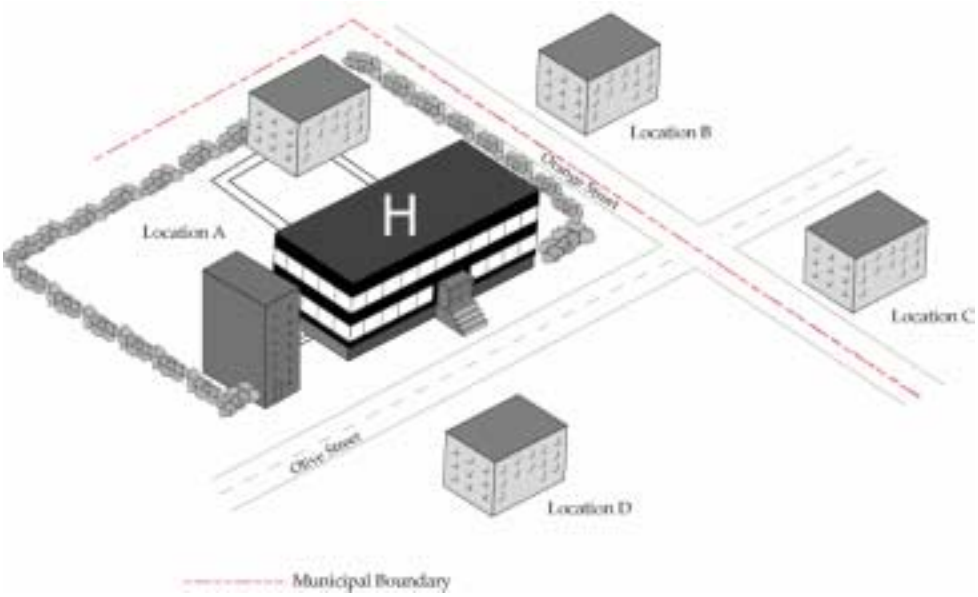

Figure 2. Sites are physical locations that provide healthcare services.

The geographic location of a site is relatively stable, as addresses most often remain unchanged. Sometimes however, an address can be modified based on changes to the healthcare service infrastructure or changes in municipalities' street naming standards.

Licenses

Healthcare services are licensed by the California Department of Health Services. These licenses contain information such as licensee, effective and expiration dates, the names and types of authorized services, as well as the address of where those services are (see Figure 5) provided.

There are a large number of licenses, in part, because of the broad array of healthcare services that require licensing and because licenses must be renewed on an annual basis.

Each license lists one or more locations (sites) where specific types of healthcare services are performed. The address information listed on the license is the primary source for the geographic location of the facilities (sites) covered by a license.

In most cases, a license authorizes a service at a single site (see Figure 3, License #2). However, a license can authorize services at more than one site (such as

"License #1). There is also not exclusivity with respect to licensed locations; some sites can provide services authorized by more than one license.

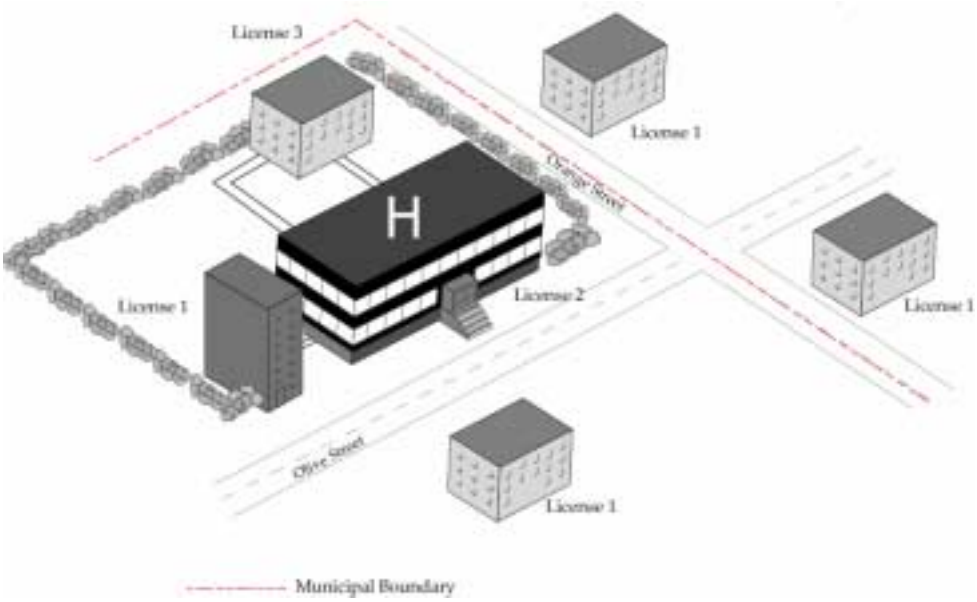

Figure 3. Licenses can be distributed across one or more sites.

Facility Identifier

Upon receipt of a license, OSHPD assigns identification codes to the primary sites covered by the license. Each identification code refers to one site (physical location). Therefore, a license can contain several OSHPD facility identifications that refer to more than one site, see Figure 4.

The criteria by which OSHPD selects sites to assign facility identifiers is not entirely clear or consistent. In those somewhat rare occasions when facility identifiers are not assigned to sites, it should be noted the sites still do have an association with a license. However, they are anomalies in the identification scheme and as such create additional complexity. Figure 4 depicts these sites as "unassigned facilities".

Figure 4. Healthcare facilities identified by OSHPD.

Summary Logical Concepts

- A ***site*** is defined as the physical location where a healthcare service is provided.
- The origin of a physical location (***site***) is an address listed on a ***license.***
- A ***license*** may contain reference to more than one address, each defining the physical location of a ***site.***
- A ***site*** is not exclusive to a particular license; its address may appear on more than one ***license*** because more than one type of service may be authorized at that ***site***.
- OSHPD assigns (in nearly all cases) a unique ***facility identifier*** to each healthcare facility listed on a ***license***.
- Not all ***sites*** listed on a ***license*** are assigned ***facility identifiers***.
- There is not always a unique relationship between a ***license*** and some other spatial context, such as municipal or zip code boundaries. This is because a ***license*** can refer to more than one ***site***.
- Because there is a relationship between a ***site*** and a ***license***, it is possible to resolve the unique combination of services that are offered within other spatial contexts, such as municipal or zip code boundaries.

SECTION 4—CONCEPTUAL DATA MODEL DESIGN

Overview

The conceptual data model is based on the three distinct subjects previously described in Section 3: healthcare licenses are represented as *licenses*, physical locations as *sites*, and OSHPD facilities as *facility identifiers*. The fundamental logical relationship between these three entities is described using an Entity Relationship Diagram (see Figure 5).

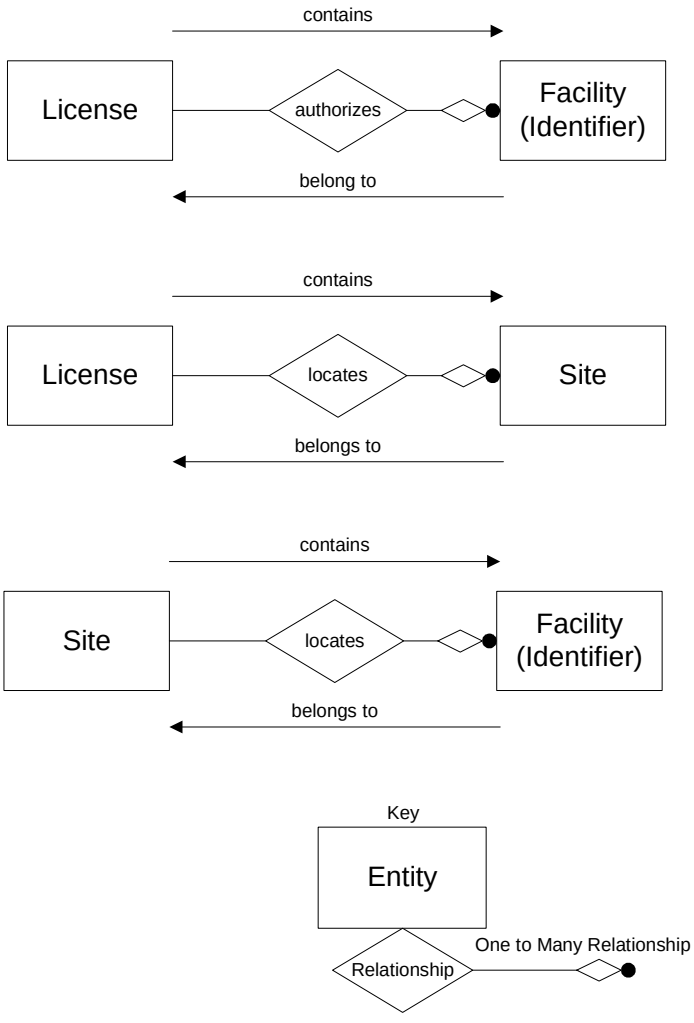

Figure 5. Entity Relationship Diagram - Healthcare Facilities Location Model.

The "real world" modeling of the three subjects is both spatial and logical. The geodatabase provides the spatial component and specifications for each individual subject, as well as the relationships between these subjects.[1,2]

This is accomplished through the following geodatabase objects:

A **Feature Class** is a collection of individual features with geometry such as points, lines, or polygons. Individual subjects (such as licenses) are designed and modeled as feature classes.

A **Relationship Class** contains a collection of relationships between feature classes and other objects as well, using pre-defined definitions. An example is the relationship between a license and the single or many addressed locations (sites) that can be covered by the license.

A **Feature Dataset** is a collection of feature classes and relationship classes that share a common spatial reference. Because all of the OSHPD Facilities occur in the State of California, all features share a common spatial reference.

Assumptions

As discussed above, it is recognized that at the present time, not all addresses found on a license are given an OSHPD facility identifier. Because the number of facilities without an OSHPD identifier is small, it is assumed that OSHPD identifiers will be assigned to those locations (sites).

The data model discussed below is based on the assumption that all facilities that are listed on a license by address and type of service will be assigned a facility identifier by OSHPD.

An alternative data model that separates (but incorporates) facilities lacking an OSHPD facility identifier is also presented, including a migration strategy if later those addresses are retroactively assigned OSHPD facility identifiers.

---

[1] Zeiler, Michael. 1999. Modeling Our World – The ESRI Guide to Geodatabase Design. ESRI Press, Redlands, CA.

[2] MacDonald, Andrew. 2001. Building a Geodatabase. ESRI Press, Redlands, CA.

Feature Dataset Class

The data model contains one feature dataset class named *HealthCareFacility*. All feature classes in this geodatabase model are contained within the *HealthCareFacility* dataset.

The dataset is defined as having a single coordinate system and all features within the dataset must use this coordinate system. Albers is often used because it provides good results with California-wide data sets. Many other statewide GIS datasets are maintained using the Albers coordinate system as well.

Another important property of the *HealthCareFacility* Dataset is establishing of X/Y Domain value ranges – the minimum and maximum coordinate value along with its precision. "Domains" specify the full range of values for a given attribute of an object. The X/Y Domain of this feature dataset must minimally cover the entire State of California. The coordinates of any features in the feature classes must fall within the X/Y domain. An example of these settings is shown in Figure 6.

Figure 6. Some example properties of the *HealthCareFacility* Feature Dataset, as seen in ArcCatalog.

Feature Classes

Three entities are modeled as Feature Classes within the *HealthCareFacility* Feature Dataset, as shown in Figure 7.

Name	Type
FacilityIdentifier	Personal Geodatabase Feature Class
license	Personal Geodatabase Feature Class
Site	Personal Geodatabase Feature Class

Figure 7. Feature Classes as depicted in ArcCatalog. Viewing these features using ArcSDE and Microsoft Windows Server is the same, except that the database name and the current user's name are attached to the feature class name.

A ***Site*** feature is an individual location where healthcare services occur. These locations originate from addresses listed on individual licenses and are assigned geographic coordinates using geo-coding services or other means. All features within the *HealthCareFacility* feature dataset consist of points which originate from those contained within the *Site* feature class.

A SiteID field uniquely defines each feature as a primary key. Additional fields might include the address information from which the coordinate location originates.

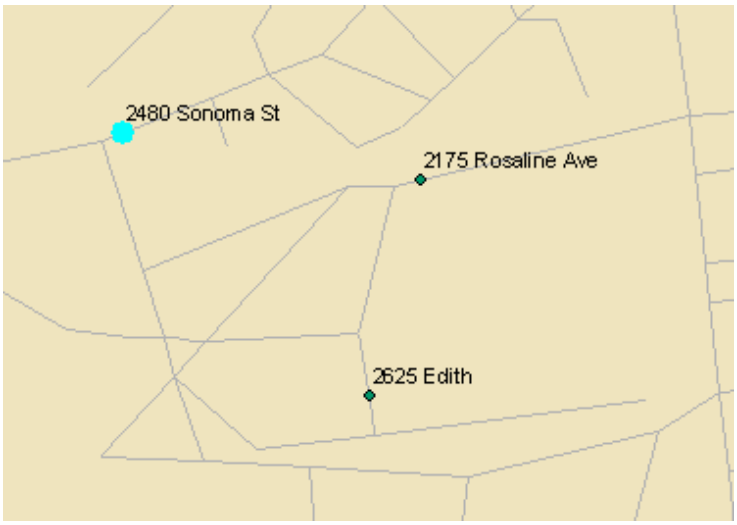

Figure 8. Three *Site* features belonging to a single license, with one site selected.

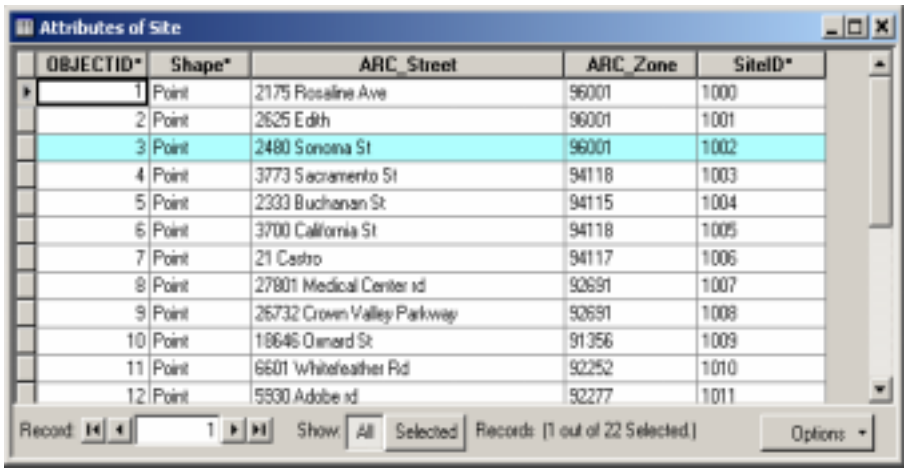

Figure 9. Attributes of the *Site* Feature Class with a single site selected.

*A **License*** feature portrays any and all locations listed on a given license. These locations must be stored as multipoint features because each license can contain more than one physical location. Each multipoint record is managed uniquely with a primary key field that stores the 9 digit numeric number assigned to each license.

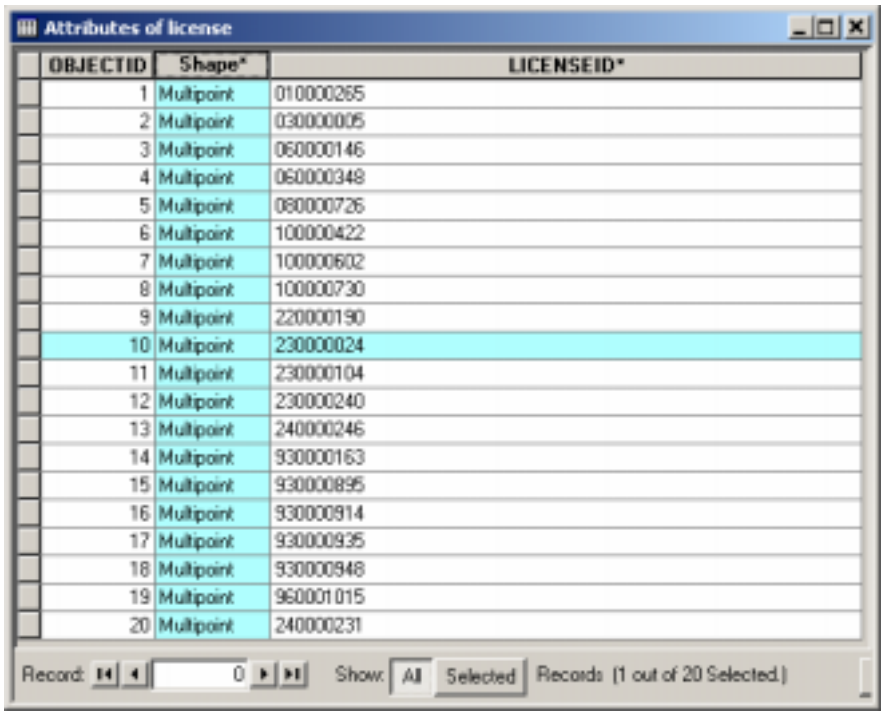

Figure 10. Attributes of the *License* Feature Class with a single license selected.

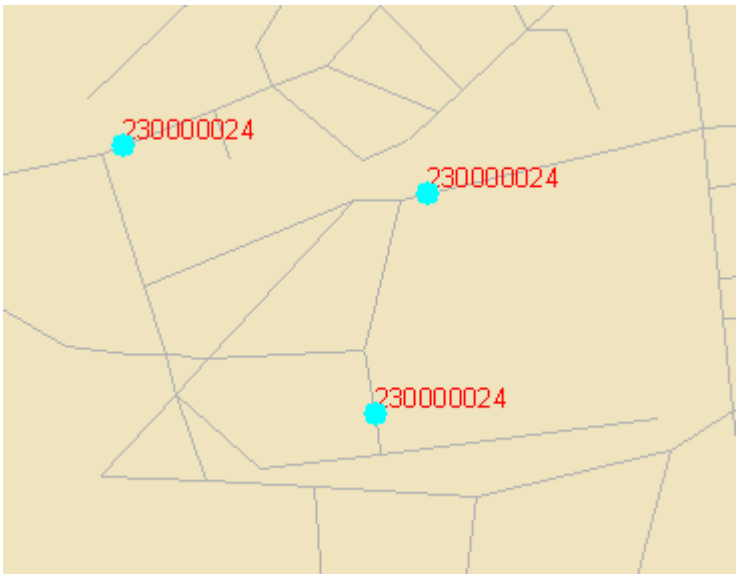

Figure 11. This single license feature selected in the *License* Feature Class contains three physical locations (sites), also depicted in Figure 8.

A ***Facility Identifier*** feature depicts a facility that is recognized by OSHPD. Since a facility only occurs at one location, it is a single point feature. The features in the *Facility Identifier* Feature Class are uniquely described using their 9 digit OSHPD identification number. Foreign key fields are also maintained in order to establish each facility's relationship with both the license that authorizes it as well as the physical site that it belongs to.

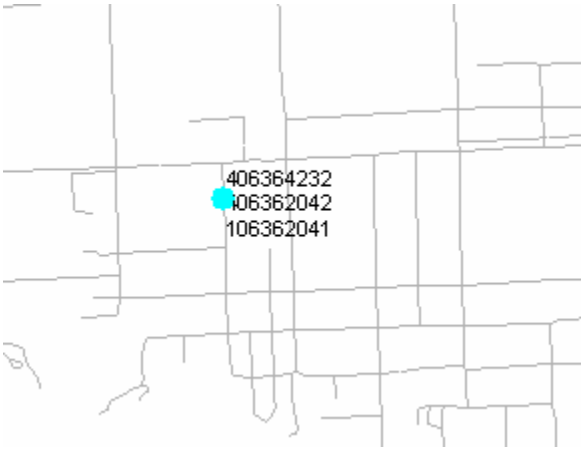

Figure 12. Three *facility identifier* features sharing one physical location (site).

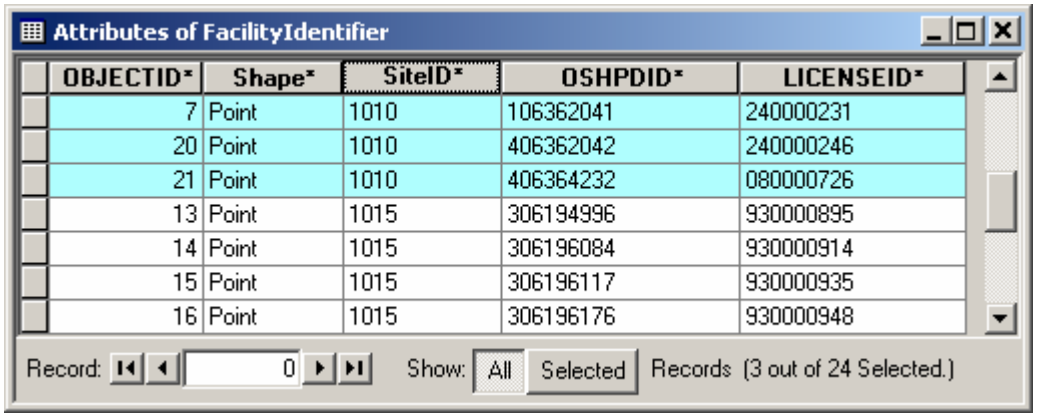

Figure 13. Attributes of the *Facility Identifier* Feature Class depict three selected features sharing one physical location (*site*).

Relationship Classes

There are two relationship classes required to establish the appropriate relationships between the three feature classes in the *HealthCareFacility* Feature Dataset, (see Figure 14).

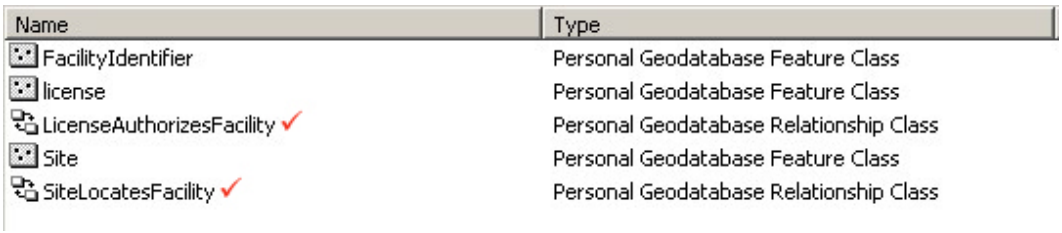

Figure 14. Healthcare Facility Dataset objects including two relationship classes.

A General explanation of relationship properties is essential in evaluating the behavior and role of these two relationships:

Name: Name of the relationship.

Origin object class: The primary feature class in a relationship.

Destination object class: The secondary feature class in a relationship.

Type: Simple relationships describe associations between data sources that exist independently of each other. If one feature is removed, the related feature in the other still class exists.

Forward Path Label: The forward path label describes the relationship when navigated from the origin to the destination. Path labels are very useful when using implemented relationships in ArcMap.

Backward Path Label: The backward path label, describes the same relationship when navigated from the destination to the origin.

Message propagation: Messaging is the mechanism for one object to notify another of a change.

Cardinality: Describes how many objects of type A are related to objects of type B.

Has attributes: Relationships can contain their own attribute values, especially between features using many-to-many types of relationships.

Origin Primary Key: Key field that uniquely identifies a specific feature.

Origin Foreign Key: Key field that uniquely identifies a specific feature from another feature class. It is another feature class's primary key.

LicenseAuthorizesFacility establishes the logical relationship between licenses and the OSHPD recognized facilities contained within them. Technical details of the relationship are listed as follows:

Name: *LicenseAuthorizesFacility*
Origin object class: *license*
Destination object class: *FacilityIdentifier*
Type: Simple
Forward Path Label: contains Facility
Backward Path Label: belongs to license
Message propagation: None
Cardinality: One to Many
Has attributes: No
Origin Primary Key: LICENSEID
Origin Foreign Key: LICENSEID

SiteLocatesFacility establishes the logical relationship between addressed locations and the facilities that are contained on the site. (Remember that a site can contain facilities authorized by more than one license.) Technical details are listed as follows:

Name: *SiteLocatesFacility*
Origin object class: *Site*
Destination object class: *FacilityIdentifier*
Type: Simple
Forward Path Label: contain Facility
Backward Path Label: belongs to Site
Message propagation: None
Cardinality: One to Many
Has attributes: No
Origin Primary Key: SiteID
Origin Foreign Key: SiteID

The relationship between *License* Features and *Site* Features is not established using a relationship class. The relationship is spatial; however the features can not be related through their data structure alone. This is because a *License* Feature is a multipoint feature and can "use" any number of *Site* Features, some of which may also be used by another *License* Feature.

The relationship between license and site can be achieved through the spatial join capability. An attribute relationship also exists in the tabular structure due to the two relationship classes, *LicenseAuthorizesFacility* and *SiteLocatesFacility*. The application software capabilities in ArcMap permit the user to associate *License* features to *Site* features through the *Facility Identifier* Feature using the two relationships.

SECTION 5—ALTERNATIVE CONCEPTUAL DATA MODEL DESIGN

Overview

Currently all addresses listed on licenses are not assigned OSHPD facility identifiers. In order to represent these service-providing locations in the data model, there needs to be a means to store these locations and maintain their relationships to the *License* features that contain them.

Additional Model Entities

Name	Type
FacilityIdentifier	Personal Geodatabase Feature Class
license	Personal Geodatabase Feature Class
LicenseAuthorizesFacility	Personal Geodatabase Relationship Class
LicenseContainsMiscSite ✓	Personal Geodatabase Relationship Class
MiscSite ✓	Personal Geodatabase Feature Class
Site	Personal Geodatabase Feature Class
SiteLocatesFacility	Personal Geodatabase Relationship Class

Figure 15. Alternative Data model components include two additional object classes.

A ***MiscSite*** feature class stores any miscellaneous addressed location that is not assigned OSHPD facility identifiers.

Figure 16. Example of a single license containing one facility and four additional addresses not assigned as OSHPD facilities - stored as *MiscSite* features.

A ***LicenseContainsMiscSite*** relationship class provides the necessary connectivity between the *License* features and any *MiscSite* features that they contain.

Name: *LicenseContainsMiscSite*
Origin object class: *license*
Destination object class: *MiscSite*
Type: Simple
Forward Path Label: Contains MiscSite
Backward Path Label: Belongs to license
Message propagation: None
Cardinality: One to Many
Has attributes: No
Origin Primary Key: LICENSEID
Origin Foreign Key: LICENSEID

Evaluation of Alternative

Maintaining a *MiscSite* Feature Class is preferable only if the licensed addresses need to be portrayed in the data model but cannot be assigned an OSHPD identification.

Maintaining the *MiscSite* Feature Class reduces some of the simplicity of implementing the data model in ArcMap for the purposes of mapping and analysis, since spatial operations will need to be performed on more than one data set.

If possible, OSHPD should assign facility identifications to all locations / services depicted on a license. Having all locations / services identified as facilities will allow more efficient GIS operations and may also reduce the potential for analytical errors associated with spatial operations.

If the *MiscSite* Feature Class *is* used and maintained, migrating those locations into the *Facility Identifier* Feature Class at a later data will not pose any logical or procedural problems as long as the locations are assigned identification numbers that are not used already.

SECTION 6— USING THE HEALTHCARE FACILITY LOCATION DATA MODEL WITH OSHPD DATABASES

The Facilities Location Data Model identifies the geographic location of the three important entities: Licenses, Facilities, and individual Sites. However, the design of these features does not account for any descriptive attribute information beyond that which is needed to establish the relationship between the three features.

To put the model to work for OSHPD GIS activities, the model must be associated with other OSHPD databases that contain descriptive attributes.

The current data model can be implemented with the LFS Mainframe Database which stores information on each OSHPD-identified facility. The LFS Data Table contains a wealth of descriptive information. Because the OSHPD identification provides the unique key to this table, its logical implementation is straight-forward with the Facility feature class, as shown in Figure 17. Examples are presented in Figure 18 through Figure 20.

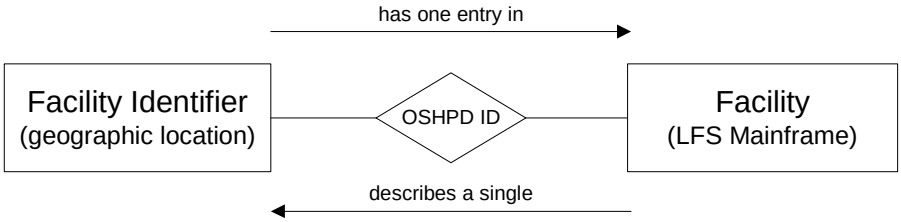

Figure 17. Simple relationship between Facility Identifier and OSHPD's LFS Mainframe Database.

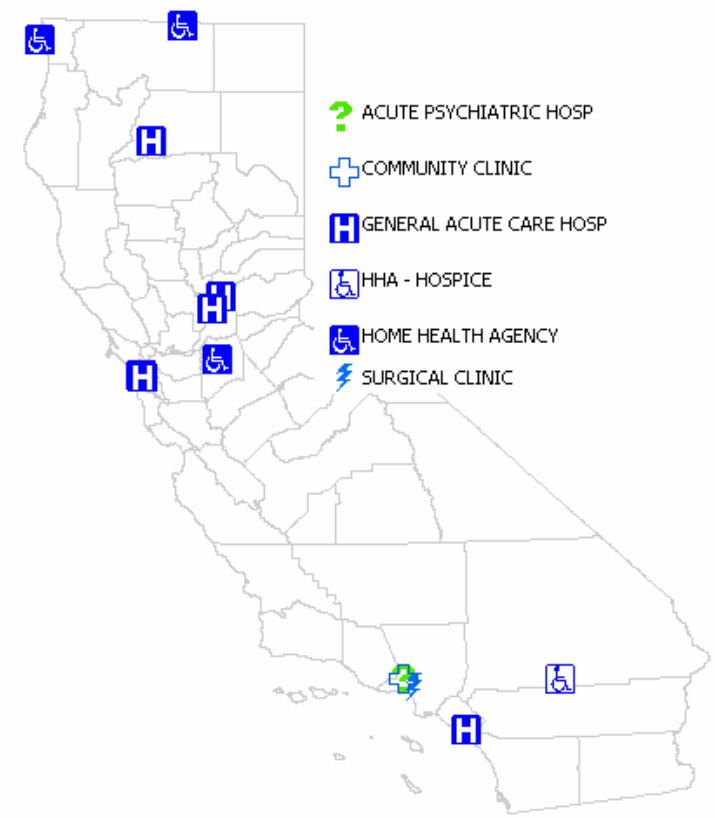

Figure 18. Facility Locations - Mapped “by Type” as derived from the LFS Mainframe Table.

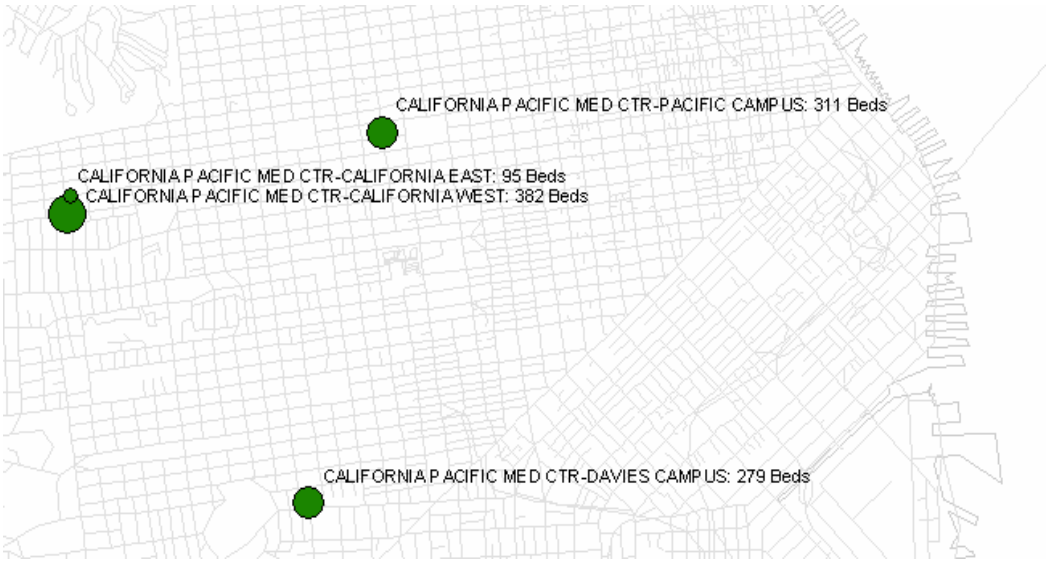

Figure 19. Name of Facility and Number of Beds, Derived from LFS Mainframe Table.

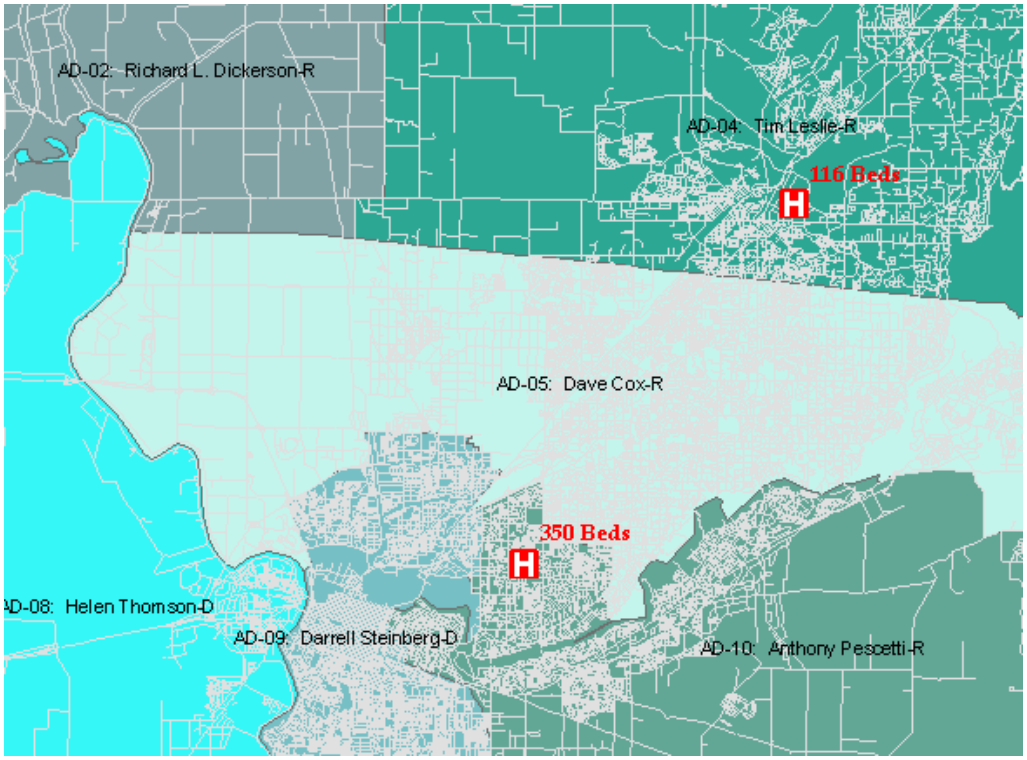

Figure 20. Hospital (and bed count) displayed with Assembly District Information.

Additional non-spatial databases are being developed by OSHPD to store healthcare information. The Medical Information Reporting for California (MirCAL) project aims to improve patient discharge data from collection through dissemination from healthcare facilities throughout the State. This includes new information from emergency departments and surgery clinics.

The Automated License Information Reporting and Tracking System (ALIRTS) is a database project under development that is re-engineering and streamlining OSHPD's business process to shift emphasis from data collection to analysis and dissemination.

A critical business requirement for analysis and dissemination of both of these projects is the incorporation of the location component of healthcare facilities.

These databases can be implemented spatially by establishing relationships between their records and the location-based structure of the spatial data model. Similar to the LFS Mainframe Database example, establishing relationships between data and location is accomplished using identifiers such as the Facility's identification number or the license's unique identification.

ADDITIONAL CONSIDERATIONS

The geodatabase objects represented in the Healthcare Facilities Location Data Model represent the minimum necessary to nominally model the relationships that exist between a number of key OSHPD databases, and the geographic representation of those facilities.

Depending on the evolving analysis needs at OSHPD, additional relationship classes could be developed to make information more readily available to the user. At this point in the data model development, it was decided to establish the "essential" elements, and that additional relationship classes could easily be established at a later date.

Specific feature class field definitions (properties, domain ranges and validation rules, etc.) have likewise not been explicitly defined at this point.

This conceptual model accommodates all the issues and variability observed through a small handful of example licenses. Future activities should include testing with a full set of California's licenses, presumably derived from the LFS Mainframe Database.

The connectivity to the LFS Database was simulated using desktop database software, and some modifications were required to obtain good results. A future activity should be to determine the requirements for defining relationship classes between data model components (such as *FacilityIdentifier*) and database objects external to the geodatabase. The "out of the box" functionality does not permit this directly; therefore direct relationships will need to be accomplished through custom development efforts.

Preliminary data model development was accomplished using ArcInfo 8.2 and a personal geodatabase. Equivalent development strategies and behaviors should be tested in a true RDMS environment such as SQL Server / ArcSDE.