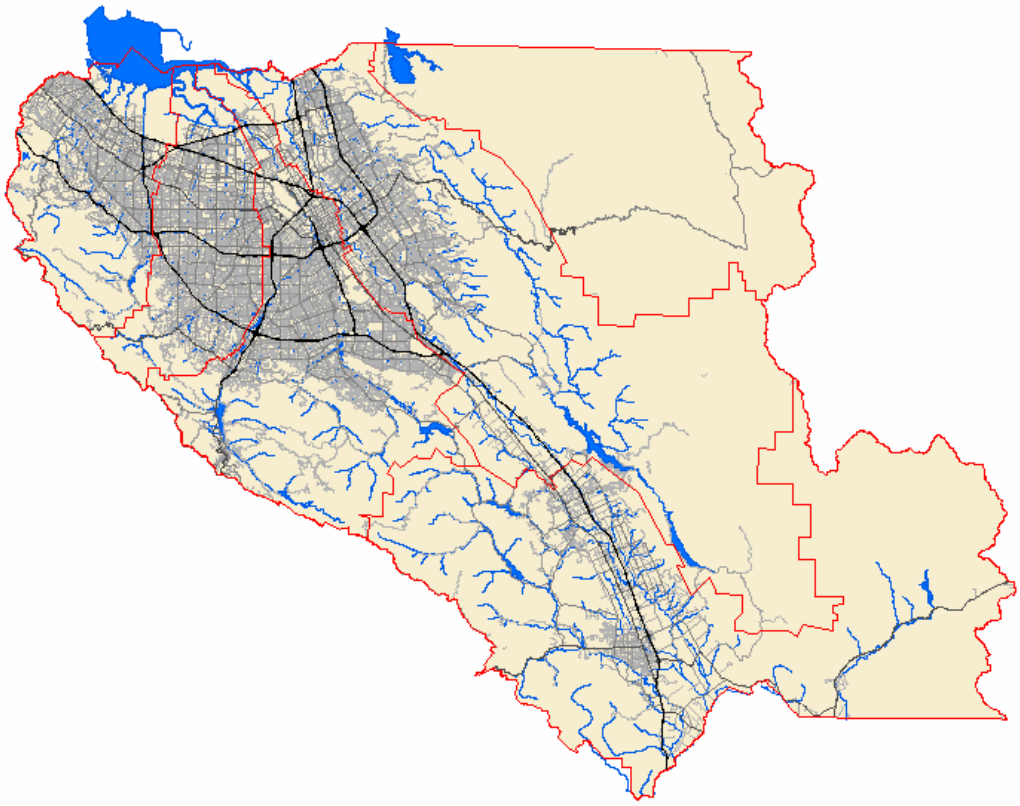
# Final Report

# Metadata Development Using The Federal Geographic Data Committee Metadata Standard in Support of the Santa Clara Valley Water District Geographic Information Systems Data

by

Ricardo Rodriguez

GIS Administration Unit
Santa Clara Valley Water District
5750 Almaden Expressway
San Jose, CA 95118
(408) 265-2600



Santa Clara County, California with watershed boundaries, major roads, creeks and reservoirs represented.

**TABLE OF CONTENTS**

## I. ABSTRACT

The Santa Clara Valley Water District's GIS data holdings were not well documented. Preliminary research on the District's GIS datasets found duplication of data, data inaccuracy, loss of information, and insufficient data documentation. Of the 572 GIS datasets, only a small percentage had documentation describing the purpose, accuracy, or creator. This problem was further compounded by the continual creation of new and undocumented GIS data sets.

The aim of this project was to implement the Federal Geographic Data Committee (FGDC) Metadata Standard within the Santa Clara Valley Water District (District). The objective was to implement a standard where GIS data would be documented, shared, and maintained as a capital resource. In doing so, tasks were created in order to successfully complete the project. The tasks at the beginning were created to allow for the examination and organization section of the project. Whereas the tasks at the end were created to allow for the communication and finalization section of the project.

The outcome of the project resulted in successfully documenting 233 District GIS datasets using the FGDC metadata standard. Additionally, the project was able to successfully meet its objectives and prove the following hypothesis: "Requiring District staff to create metadata for GIS datasets before placing GIS datasets on the GIS server will reduce the amount of undocumented and unneeded GIS datasets."

## II. INTRODUCTION

Santa Clara County is situated between the Santa Cruz Mountains and the Diablo Mountain Range. Santa Clara shares borders with Santa Cruz and San Mateo counties to the west, Merced and Stanislaus counties to the east, San Benito County to the south, and Alameda County to it's north. Located south of San Francisco Bay and North of the City of Gilroy, Santa Clara County is one of California's busiest urban areas.

In the early 1920's, Santa Clara County was composed mostly of farmers and ranchers who held large parcels of land in both urban and rural areas. Orchards and vineyards grew in large numbers, and a variety of vegetables saw cultivation as the century drew to its close (ref 1). From what had been a predominately agriculture economy in the 1920's, Santa Clara County became a combination of industrial-agricultural economy in the 1940's. The county's economy had grown into a market and shipping center with a related network of food processing, food packaging, and manufacturing of food machinery. Following the 1950's, Santa Clara County had seen more and more industries such as Ford Motor Company, Lockheed, and General Electric move into the County (ref 1). Today, technology companies such as Oracle, Google, Cisco, and Apple have begun to dominate the county's landscape.

As more and more industries moved into Santa Clara County, the population increased. In the decade from 1940-1950, the growth in Santa Clara County's population increased by 66% to 291,000, and in 1950-1960 an increase of 121% to 642,000 took place (ref 1). Currently, the county's population is estimated at more than 1.7 million, with cities such as Los Altos Hills,

Gilroy, and Milpitas recording 2% or more growth from 2001-2002. Overall, Santa Clara County ranks as the fifth highest populated county in California, and its population is estimated to reach 2.2 million by 2020 (ref 2).

## III. BACKGROUND

The District is located in Santa Clara County and resides within the boundaries of the City of San Jose. The District's history can be traced to the 1920's when approximately fifty farmers met to formalize the establishment of the Santa Clara Valley Water Conservation Association to work for water conservation through more efficient agricultural methods (ref 1). In 1951, legislation was enacted which created a second agency known as the Santa Clara County Flood Control and Water Conservation District whose primary responsibility was flood control. Over the years, conflicts arose within the two agencies as they struggled for political control to determine who should control and manage Santa Clara County's vanishing water supply. Concerns over the duality of the two Water Districts ultimately resulted in the merger of the two agencies and established the Santa Clara County Flood Control and Water District in 1968. In January 1, 1974, the name changed once more to the current name Santa Clara Valley Water District (ref 1). The role of the District has not changed over the years. Along with flood protection and water conservation, the District works to provide Santa Clara County with enough clean and safe water for homes and businesses. As the primary water resource agency in the County, the District not only acts as the water wholesaler, but also as the steward for its watersheds, streams and creeks, underground aquifers, and district built reservoirs.

To improve on the services provided to the residents of Santa Clara County, the District began using Geographic Information Systems (GIS) technology. Over the past 14 years, a significant amount of time and money has been invested in GIS technology at the District. First implemented as a Pilot Well Head Protection Project, GIS was used to gather information about hazardous material sites in order to protect wells from contamination. With the addition of Santa Clara County parcel data, GIS was used to determine if parcels were at risk of flooding. Currently, GIS is being used in many District projects to analyze and query geographic features such as creeks, pipelines, and District owned land in order to present information to managers, staff, the public, and Governing Board members.

The existing GIS technology has provided significant value to many users within the District's Divisions and Units. District staff have benefited in using GIS for data development, mapping locations, and presenting information. By industry standards, the sheer number of District GIS users represents the tremendous success of GIS within the District. However, although useful, the GIS Administration Unit within the District believes the GIS investments are not yielding their full potential benefits, and can only be realized through a broader, more coordinated GIS environment.

## IV. OBJECTIVES

The GIS Administration Unit seeks to maximize the benefits of the District's investment in GIS. The mission of the GIS Administration is to develop and maintain a District Enterprise Geographic Information System infrastructure and process framework that enables District staff to effectively utilize Geographic Information Systems technology. As so, the GIS Administration Unit has determined there are a number of opportunities to improve the management and use of GIS technology for both the effectiveness and efficiency of District Divisions and Units in carrying out their missions.

One major opportunity in improving the District's GIS is through management of GIS data. The District's existing GIS data holdings are not well documented. Preliminary research on GIS data has found duplication of data, data conflicts, data inaccuracy, loss of information, and insufficient data documentation.

In this project, we will implement the FGDC Content Standard for Digital Geospatial Metadata within the District. The FGDC Content Standard for Digital Geospatial Metadata (metadata) promotes coordinated development, documentation, sharing, and dissemination of geographic data (ref 3). This project addresses the problems related to geographic data documentation, data inaccuracy, loss of information, duplication of data, and insufficient metadata within the District. The overall project objectives are:

1) Implement the FGDC Metadata standard within the District for proper documentation of GIS datasets.
2) Assist District staff on creating proper metadata documentation for GIS datasets.
3) Make available and distribute a revised list of GIS datasets with accompanying FGDC Metadata documentation.

Accomplishing the above objectives will result in an improved process for documenting and distributing GIS data. Overall benefits to the District will be improved data for decision making, analysis, and sharing of information. This project will also lay the foundation for the building of an Enterprise GIS system within the District.

## V. HYPOTHESIS

The following hypothesis explains in more detail how implementing the FGDC Metadata Standard will help improve the District's GIS data.

*Hypothesis:*
*Requiring District staff to create metadata for GIS datasets before placing GIS datasets on the GIS server will reduce the amount of undocumented and unneeded GIS datasets.*

The GIS Administration controls which GIS Datasets are placed on the GIS server and made available to District staff. If District staff would like a GIS dataset placed on the GIS server, the following process must take place: 1) District staff must contact the GIS Administration and ask

that a GIS dataset be placed on the GIS server, 2) District staff must provide the GIS Administration with the GIS dataset, and 3) the GIS Administration will place the GIS dataset on the GIS server for District distribution. District staff can then access the dataset by navigating through the GIS server folder structure, or by using a GIS extension within a GIS desktop software program. Although this process is effective for distributing GIS data, it also contributes to the District's current GIS data problems. The current process inadvertently allows for the buildup of GIS Datasets that are out dated, inaccurate, or lacking documentation describing the purpose or creator of the data.

Metadata, also known as "data about data" or "information about data", describes the context of GIS datasets such as who, what, when, why, and where the GIS dataset was created. District GIS users are currently not required to complete Metadata in order to have their datasets available for distribution on the GIS server. By changing the current procedure, and requiring District staff to create and provide metadata with their GIS dataset, we will reduce the amount of undocumented and unneeded GIS datasets.

In keeping with the FGDC, this project will document GIS data, teach metadata standard documentation, and allow data to be shared internally and on the National Geospatial Data Clearinghouse. Furthermore, this project will allow for more than adequate documentation of the District's current and future GIS data holdings.

*Prediction:*
A number of District staff will be unenthusiastic in creating and providing metadata for GIS datasets, therefore, they will not request GIS datasets to be placed on the GIS server. In contrast, a number of District staff will take the time to create metadata for their GIS datasets to be placed on the GIS server, therefore, contributing in enhancing the District's GIS database.

## VI. METHOD

The following tasks define the method that was implemented in this project. Each of the tasks describe the major elements that were completed to achieve the stated objectives.

**Task 1 – Develop Project Team**
The role of the GIS Administration is to develop and maintain a District Enterprise Geographic Information System within the District. As so, the team for this project consisted of GIS Administration Unit staff members Kurt Hassy and Ricardo Rodriguez. The purpose of a two-member project team was to minimize delays caused by conflicting views, differing approaches, or personality differences that occur within larger project teams. The project team was also created to efficiently and effectively utilize the team's talents. The talents of Kurt Hassy in GIS software applications, metadata creation, and application development helped in analyzing GIS data, creating metadata, and developing applications. The experience of Ricardo Rodriguez in GIS software applications, database management integration, and metadata creation was used in analyzing and organizing GIS data, teaching and creating metadata, and completing the final report.

**Task 2 – GIS Data Collection and Inventory**
Previous GIS data inventory work completed in the year 2000 was the foundation for task 2. The GIS data inventory work consisted of a spreadsheet listing of all known GIS datasets within the District. The organization of the GIS datasets were placed into 11 categories: biology, creeks and reservoirs, district facilities, flooding, groundwater management, map grids and miscellaneous, parcels, physical geography, political boundaries, roads and transportation, and watersheds. The GIS data inventory also contained relevant information about GIS data such as data location, last data update, data type, and assumed data owner. Although useful, the GIS data inventory was infrequently updated.

To have a better understanding of the current GIS datasets within the District, the GIS data inventory spreadsheet was updated. The GIS data on the District's two servers were analyzed and compared to the GIS data inventory spreadsheet. Work for this task was divided between the two project members. Ricardo Rodriguez assumed responsibility in updating the GIS data in the following categories: biology, creeks and reservoirs, parcels, physical geography, and watersheds. Kurt Hassy assumed responsibility in updating the remaining categories: district facilities, flooding, groundwater management, map grids and miscellaneous, political boundaries, and roads and transportation. The GIS data that was not previously documented in the GIS data inventory spreadsheet was promptly added to the inventory.

**Task 3 – Determine Owners of GIS Data**
After the GIS data inventory spreadsheet was updated, we located the District staff responsible for the data. Given that the GIS Data inventory spreadsheet had listings of assumed owners of GIS datasets, we decided to contact these owners first. For datasets with no owners associated to them, we developed a method to determine who to contact. The method was to overlay questionable GIS data layers over Santa Clara County watershed boundaries. Once the questionable datasets were added, they would fall within the boundaries of certain watersheds. Because the District has various business units organized by watersheds, we formalized that data within a certain watershed will most likely belong to business units associated with those watersheds. Implementing this non-scientific method, we were able to locate owners for the majority of questionable GIS datasets.

**Task 4 - Contact and Distribute Email Information Package**
The decision was made by the project team to send an email (appendix A) to the Unit managers of individuals who were initially determined to be the owners of GIS data. This decision was made in order to inform the Unit Manager of the ongoing project, and provide him/her with instructions on what to do and who to contact concerning any questions. This was also done in an effort to ensure that assumed owners of GIS data would receive the email information directly from their managers. The expectation was this would reduce the likelihood of people avoiding participation in the project.

The email sent to the Unit Managers also contained an email attachment named Meta_Package.zip, which contained three documents: 1) a memo (appendix B); 2) a metadata template document (appendix C); and 3) a metadata help document (appendix D). The memo informed the Unit manager of what additional information was needed, a list of his/her Unit staff

members who were presumed to be the owners of GIS datasets, and a list of GIS datasets. The metadata template document contained the necessary descriptive data fields required for GIS data to be compliant with FGDC metadata standards. The metadata help document provided information describing and explaining the descriptive data fields in the metadata template document.

We requested the Unit Manager to direct his/her staff to examine the GIS datasets to determine whether they owned the data. In case his/her staff owned the GIS data, and wanted the GIS data placed on the server, we requested they complete the metadata template document and return the Meta_Package.zip. In the likelihood they were not the owners of the GIS data, or declined to complete the metadata template, they were excluded from the project.

**Task 5 – Collect GIS Packets and Convert Metadata**

Once the Meta_Package.zip packets were collected, the metadata template document within the Meta_Package.zip was converted into an electronic metadata file. A Metadata Transfer application was developed to convert the metadata template document information into the .xml required programming language used by the FGDC- Environmental Systems Research Institute metadata style sheet (appendix E). Once all metadata was converted, QA/QC was performed on the FGDC-ESRI metadata style-sheet to correct any errors or missing information that may have occurred during the metadata transfer process.

**Task 6 – Determine Which Remaining GIS Datasets to Keep**

In the final task of the project, we determined which GIS datasets would be placed on the GIS server. As stated in the memo sent in the Meta_Package.zip, all GIS datasets returned with completed metadata files were placed on the GIS server. To determine the outcome of GIS datasets returned with no metadata, we requested help from the GIS Administration Unit staff.

The GIS Administration Unit staff and project team agreed to review and analyze the remaining GIS datasets before deciding to archive or place them on the server. Two to three hour meetings were scheduled once a week for one month to complete the review of the GIS datasets. Meetings consisted of a laptop running a GIS desktop program, and a laptop projector that displayed the questionable GIS datasets for review. With the familiarization of the GIS datasets by the GIS Administration Unit staff, and the use of existing GIS datasets for comparison, we were able to determine which GIS datasets to place on the server.

After determining which GIS datasets to keep, the issue remained of creating metadata. The decision was made to distribute the remaining GIS datasets between the GIS Administration Unit staff and the project team to create Metadata for the remaining GIS datasets.

## VII. RESULTS

As described in Section VI (Method), the project team was able to successfully complete this project. This was obtained through the help of many individuals within the District, most notably, the assistance of the GIS Administration Unit staff. The tables below show the results of the overall completed project.

*Task 2 - GIS Data Collection and Inventory*

The amount of GIS datasets at the start of the project was 561. After searching and analyzing the two GIS servers, the amount of GIS datasets grew to 572 (Table 1). The growth in GIS datasets resulted from new GIS datasets being added to the GIS servers, and GIS datasets being overlooked during the initial tally of GIS datasets. Observing Table 1, one can notice a decline in some of the "# of GIS Datasets" in the "GIS Datasets Updated" column. In these situations, the decline was caused from the project team finding and removing duplicate GIS datasets, or from GIS datasets being lost.

**Table 1. Represents the number of GIS datasets at start of project and the new number of GIS datasets after GIS servers were analyzed and reviewed.**

| **GIS Datasets at Start of Project** | | **GIS Datasets Updated** (After GIS Servers Analyzed and Reviewed) | |
|---|---|---|---|
| ***GIS Folder*** | ***# of GIS Datasets*** | ***GIS Folder*** | ***# of GIS Datasets*** |
| Biology | 62 | Biology | 55 |
| Creeks & Reservoirs | 97 | Creeks & Reservoirs | 98 |
| District Facilities | 19 | District Facilities | 21 |
| Flooding | 39 | Flooding | 40 |
| Groundwater Management | 61 | Groundwater Management | 61 |
| Map Grids & Misc. | 29 | Map Grids & Misc. | 32 |
| Parcels | 104 | Parcels | 109 |
| Physical Geography | 23 | Physical Geography | 25 |
| Political Boundaries | 83 | Political Boundaries | 86 |
| Roads & Transportation | 32 | Roads & Transportation | 33 |
| Watersheds | 12 | Watersheds | 12 |
| **Total** | **561** | **Total** | **572** |

*Task 4 – Collect GIS Packets and Convert Metadata*

The number of GIS datasets emailed to prospective owners was 572, and the number of GIS datasets returned with completed metadata was fifty-four (Table 2). The nominal return of GIS datasets supported our hypothesis for this project:

*Hypothesis:*
*Requiring District staff to create and provide metadata for GIS datasets before placing GIS datasets on the GIS server will reduce the amount of undocumented and unneeded GIS datasets.*

There were several reasons for the nominal return of GIS datasets with metadata. One reason was individuals declining to create metadata because of time issues. Through personal communication, many individuals stated they did not have time to create metadata, however, they deemed their GIS datasets were important. The second reason was GIS datasets were emailed to the wrong owners. This was a concern at the beginning of the project, and was

anticipated to affect the outcome of the number of metadata files returned. Due to insufficient GIS dataset information at the beginning of the project, we could only use our best judgment, personal communication, and existing information to determine the owners. Additionally, there may have been individuals who did not claim GIS datasets in order to avoid creating metadata.

**Table 2. Represents the number of GIS datasets that were emailed to perspective owners and the number of GIS datasets that were returned with appropriate Metadata documentation.**

| **GIS Datasets Emailed Out** | | **GIS Datasets Returned** | |
|---|---|---|---|
| ***GIS Folder*** | ***# of GIS Datasets*** | ***GIS Folder*** | ***# of GIS Datasets*** |
| Biology | 55 | Biology | 39 |
| Creeks & Reservoirs | 98 | Creeks & Reservoirs | 9 |
| District Facilities | 21 | District Facilities | 1 |
| Flooding | 40 | Flooding | 1 |
| Groundwater Management | 61 | Groundwater Management | 2 |
| Map Grids & Misc. | 32 | Map Grids & Misc. | 0 |
| Parcels | 109 | Parcels | 0 |
| Physical Geography | 25 | Physical Geography | 2 |
| Political Boundaries | 86 | Political Boundaries | 0 |
| Roads & Transportation | 33 | Roads & Transportation | 0 |
| Watersheds | 12 | Watersheds | 0 |
| **Total** | **572** | **Total** | **54** |

*Task 6 – Determine Which Remaining GIS Datasets to Keep*

The amount of GIS datasets returned with completed metadata was fifty-four, and the amount of GIS datasets that were not returned with completed metadata was 518 (Table 3). The nominal amount of GIS datasets returned was expected at the start of the project. The GIS Administration Unit and the project team dealt with the remaining GIS datasets as described in Task 6 in Section V – Method.

**Table 3. Illustrates the number of GIS datasets that were returned with appropriate Metadata documentation and the number of GIS datasets that were not returned.**

| **GIS Datasets Returned** | | **GIS Datasets Not Returned** | |
|---|---|---|---|
| ***GIS Folder*** | ***# of GIS Datasets*** | ***GIS Folder*** | ***# of GIS Datasets*** |
| Biology | 39 | Biology | 16 |
| Creeks & Reservoirs | 9 | Creeks & Reservoirs | 89 |
| District Facilities | 1 | District Facilities | 20 |
| Flooding | 1 | Flooding | 39 |
| Groundwater Management | 2 | Groundwater Management | 59 |
| Map Grids & Misc. | 0 | Map Grids & Misc. | 32 |

| | | | |
|---|---|---|---|
| Parcels | 0 | Parcels | 109 |
| Physical Geography | 2 | Physical Geography | 23 |
| Political Boundaries | 0 | Political Boundaries | 86 |
| Roads & Transportation | 0 | Roads & Transportation | 33 |
| Watersheds | 0 | Watersheds | 12 |
| **Total** | **54** | **Total** | **518** |

The project team and the GIS Administration Unit decided to create metadata for 179 of the 518 GIS datasets. Fifty-four GIS datasets were returned with completed metadata from District staff. In total, 233 GIS Datasets were placed on the GIS server (Table 4).

**Table 4. Represents the number of GIS datasets returned with associated metadata and the number of GIS datasets retained.**

| GIS Datasets Returned | GIS Datasets Retained | Total |
|---|---|---|
| **54** | **179** | **233** |

## VIII. CONCLUSION

At the beginning of this project, the District had very little documentation for their GIS data. The documentation that did exist for GIS data was either missing critical information or incorrect. In order to correct this problem, we implemented the FGDC Content Standard for Digital Geospatial Metadata within the District.

The expectations we had for this project were to meet our objectives in section IV – Objectives, and to prove our stated hypothesis that "*Requiring District staff to create metadata for GIS datasets before placing GIS datasets on the GIS server will reduce the amount of undocumented and unneeded GIS datasets.*" I believe we met the objectives and proved the hypothesis of this project, as explained in section VI – Method and section VII - Results of this paper. Although the objectives and hypothesis for this project were met, there were difficulties we encountered.

Many of the difficulties we encountered, such as District staff not participating, were predicted at the beginning of the project. As a result, we were able to accept this problem and not get discouraged during the project. Many of the meetings and discussions we had with District staff proceeded as expected. Other problems, such as converting the metadata documentation template into .xml language, were not predicted. In this situation, we expended more time finding a solution for the problem because we did not predict it early on in the project. Fortunately, we were able to predict the majority of problems that arouse during this project, and therefore formulate proper solutions.

Ultimately, we were able to successfully implement the FGDC Content Standard for Digital Geospatial Metadata within the District. We were able to create metadata for 233 out of 572

District GIS datasets. We were able to archive 339 GIS datasets that were poorly documented or unneeded. Most importantly, we were able to inventory and preserve the District's investment in GIS data, and create a standard operating practice for creating metadata within the District.

## IX. REFERENCES


1. California History Center De Anza College. 1981. *Water in the Santa Clara Valley: A History*. Local History Studies-Volume 27.

2. The State of California Employment Development Department WebSite (2003) Available http://www.calmis.ca.gov/ [19 August 2004]

3. Federal Geographic Data Committee. 1997. *Framework Introduction and Guide*. Federal Geographic Data Committee. Washington, DC.

## X. APPENDICES

Appendix – A : E-mail sent to District Unit Managers concerning the Metadata project.

Appendix – B : Memorandum sent as attachment with email sent to District Unit Managers.

Appendix – C : Metadata template document sent as attachment with email sent to District Unit Managers.

Appendix – D : Metadata help document sent as attachment with email sent to District Unit Managers.

Appendix – E : FGDC-ESRI metadata style-sheet used to display metadata information.

## XI. Discussion/Integration of Science and Policy

(science and policy)

At the beginning of this project, we had certain expectations and thoughts of what the result of this project would be. One expectation is that people would be eager to help and assist in the cleaning up of Santa Clara Valley Water District GIS data. Another is that it would allow us to find the individuals who update certain datasets, and also allow the implementation of a GIS Metadata Standard. However, there were a few problem which existed that we did not take into account.

(the problems exist integrating science and policy)

One problem that existed that we did not take into account was the complication of integrating the work of scientist with the policy of GIS standards. Much of the GIS data files which needed metadata was file concerning environmental data such as habitat area for red legged frogs and restoration areas did not have proper metadata. As we soon found out during this project, this was due to communication to scientists such as biologist, vegetation managers not understanding what data was needed for Metadata, or lack of knowledge of GIS.

Another problem which existed is the fact that scientist did not understand that GIS data could be used as a source for the public, and can track long term effects of areas of interest. This became clear as the project continued on throughout the task. What was notices was that scientist would use data for own project with no thought of it's use for future projects and tracking. No thought of data being used for others using the data such as managers, engineers, city planners, etc.

(GIS enables both policy and scientist to easily visually see graphical things that they can understand.)

The creation of Metadata for GIS datasets benefits scientist and policy makers. By allowing scientist to find data and find out important information. This also allows for

(how do we get scientist to integrate technological policy, so they can bridge the gap between science and policy. Ethically showing the correct information.)

Showing scientists that Metadata allows for them to gather information that shows the accuracy, who created, what , where, when, and why the data was created that is beneficial for their overall project. Once this was showed, scientist were more eager to help in creating metadata. For their project.

---

(reference some papers.)

---

(problems ran into, difficulties we had how we should have fixed them, or done different)

Other problems we ran into dealt with the environment and culture of the District. One barrier was the lack of support from the District as a whole for the use of GIS technology. This added to the problems because people felt that it was an unimportant effort to clean up GIS data. This is contributed to the lack of knowledge people have of the use of GIS and it's importance when working on projects.

A second problem dealt with having people fill out the Metadata format for their respective data set. We had thought at the beginning of the project that SCVWD staff would be more acceptable of the project and willing to help. As it turned out, people were willing to find out what we were doing, but did not , or was hesitant to claim data and fill out metadata for the data. Staff did not want to take responsibility for data sets and their updates. As a result, the GIS Administration took it upon themselves to complete metadata for the remaining GIS datasets.

(what is next step)
Working through the problems, allowed the successful completion of the project and started off in the right direction for continued growth of standards for GIS within the SCVWD. Next steps within the District are to implement a naming convention for GIS datasets which are put on the server, and create a logical hierarchal structure within the folder which allows for easy navigation and search for GIS data. This step will be completed with the help of a consultant.

The hiring of a consultant will be the next step to move the GIS technology forward within the District. The main focus of the consultant will be to perform a GIS needs assessment of current GIS environment within the District. Mainly focusing on business Unit goals, missions, and objectives. Finding out how GIS is used in their daily work flows, what gis data they use, and what is their current need in gis. The results of this project will then determine the direction in which the GIS Administration will take for the next 2 years.

## 4. Presentation of Background Information and Literature (Objectives)

[This plan will serve as the foundation for the overall goals, strategies, and specific requirements for the GIS Administration Unit in implementing a District Enterprise GIS to address the diverse needs of Divisions, Units, private contractors and the public. The following four objectives are critical to the long term success of a coordinated GIS effort to support the structure and functions of the District.

1. Institutional Focus

   Until recently, the District de-emphasized the role of a core GIS support unit. As the role of the core GIS support unit diminished and weakened, GIS coordination between Divisions and Units decreased, redundant efforts increased, and some Units were left unsupported. One lesson learned from this experience is that a coordinated effort between a core GIS support unit, Divisions and Units must exist in order to create a successful District Enterprise GIS.

   A primary objective is to establish an effective Unit with a clear mission and mandate. For the purpose of this plan, the GIS Administration Unit is deemed to be the core GIS support unit for coordination of GIS activities.

Critical to the long term success of the GIS Administration Unit is the unit structure and financial base that forms the foundation for its operations. Without a strong and consistent financial base, the GIS Administration Unit will have limited success.

2. Data Focus

The District is heavily dependent on data and information in its decision making and operations. In many ways, the District is data rich. In other ways, while there are volumes of data used and created, the data does not always meet the needs of Divisions and Units for various projects. Data issues include accuracy, accessibility, duplication , completeness, and loss of information.

A primary objective is to document the status of data and related information. The purpose is to address data from an enterprise perspective, and to coordinate data use and data development as a capital resource through Divisions and Units in a more unified and effective method.

3. Business Focus

The success of the Enterprise GIS will be determined in large measure by how well it enables the various Divisions and Units within the District to improve their functions. The Enterprise GIS must accommodate Division and Unit needs, and offer a level of enhancement over their existing tools and processes.

A primary objective is to implement GIS technology where Divisions and Units spend less time supporting their own GIS systems, and benefit from tools and data of a supported District Enterprise GIS.

4. Project Vesting

The success of an Enterprise GIS is dependent on the active participation of Divisions and Units. To ensure participation among Divisions and Units, the creation of an Executive Level GIS Steering Committee is needed. The Executive Level GIS Steering Committee is the first step in fostering coordination between District Divisions and District Units. Only if Divisions and Units have a voice in the GIS systems adoption and systems governance will they make the District Enterprise GIS successful.

A primary objective is to involve Divisions and Units in GIS by establishing an Executive Level GIS Steering Committee to promote the development of an Enterprise GIS governance structure. ]

## 5. **Rational For Project**

1. District provides stream stewardship to County

The Santa Clara Valley Water District provides flood protection, stream stewardship, reliable water supply, and flood control to more than 1.6* million residents in the Santa Clara County.

Recently county residents provided the SCVWD with the responsibilities of concentrating more effort on watershed activities and flooding control by passing the Clean, Safe Creeks and Natural Flood Protection Program initiative. The passing of this initiative allows for a number of projects aimed at providing long term flood protection, stream stewardship, flood protection, and a healthy ecosystem watershed. Projects concentrating on this issue will be implemented in the next 15 years resulting in a plethora of information and data.

2. GIS data is used to make decisions on policy, construction, projects, restoration efforts.

The information and data subsequently created by the numerous projects around the District, will generate spatial information("use quote on Govt use of spatial data"), and may also need spatial information to continue on the projects. The data subsequently will be used to further enhance District projects by being used in presentations, reports, used to support construction efforts or policy decisions. Current efforts are the construction on pipeline retrofitting which supplies water to Santa Clara County residents from the Hech Heche pipline. Other projects are restoration efforts on such creeks as the Guadalupe River to try and stop flooding. In order to make the best decisions possible on projects, GIS data must be reliable and accurate.

3. GIS data is inaccurate, people do not know who developed the data.

Reliable data is what is currently hindering the advancement of GIS and its ability to find solutions to managerial, environmental, and other problems. Currently, the District GIS data is not as accurate as it can be. There are many incidents where data is used that is outdated or inaccurate. Currently the District has four layers representing watershed boundaries, which all represent different boundaries for the same watershed. This inaccuracy is magnified if you take into account that no one knows where the data comes from, or who created the data. The watershed data set is no the only data set with inaccuracy. Other core SCVWD datasets such as the creeks, pipelines, infrastructure, and wells data are also inaccurate. However, such data is still being used to guide policy, construction, and restoration decisions.

4. Implement an FGDC standard to track who creates data, and accuracy.

To improve the current situation of the SCVWD GIS data, a Metadata standard must be implemented in order to track changes, development, editing, and sharing of information. After performing research, it was decided that the Federal Geographic Data Committee's Metadata Standard would be implemented and used in order to track current and future creation and changes of GIS data. The FGDC Metadata standard will help by providing the basis of needed information of GIS Data of who created it, when it was created, what was it used for, contact information and other information that is relevant. By implementing the FGDC Metadata, the District takes it's first step in the start to begin to provide accurate data to the district and the public.

## 6. Description of Project (Method)

The methods used for this project were based on a two person work effort. Work began on the Metadata project on September 2003 with the work effort of two individuals: Rick Rodriguez

and Kurt Hassy. A project timeline was first created which set mini-project completions for task needed in order to move the Metadata Project forward.

1. First step was to obtain a list of current SCVWD GIS data.
2. Second step was to create a metadata template / implement metadata standard to be used within the district.
3. Third step was to contact who we believed the owners of the GIS Data were.
4. Fourth step was to have meeting to go through the SCVWD GIS Data and select the priority datasets.
5. Fifth step was to organize the priority datasets and the individual owners and create metadata for GIS Data.
6. Sixth step will be to place metadata on the ESRI Metadata Explorer .
7. Seventh step will be to contact the Federal Geographic Data Committee and set up the harvesting tool to make periodic uploads to the NSDI, and share information to the public.

## 7. Presentation and Discussion of Results

The results of our effort produced an estimate of 120 GIS datasets from the start of 600 datasets. The 120 datasets were chosen as stated in the methods section of this report. The 120 datasets are the remaining sets that have sufficient metadata that complies with the Federal Geographic Data Committee. The following is a list of the remaining GIS datasets that will be transferred to the new GIS Server, and used as the core for the Santa Clara Valley Water District.

| | | | |
|---|---|---|---|
| | | | |
| | | | |

## 8. Integration

The importance of GIS data is currently being realized. The use of GIS is now currently being used in many government, state, local, and private areas. From location based services, to routing, to environmental impact analysis, remotes sensing, business analysis, GIS use and it's importance is contributing to many industries. The importance of GIS, and GIS coordination and information sharing was best seen in the aftermath of 9/11 when GIS was used to assist in finding where structures existed, and using thermal imaging used to determine where hazardous search sites were in order to avoid potential danger. With the use of GIS in so many different fields, and with the many applications for it's use, it is determined that the SCVWD can also benefit from an organized GIS system that allows for data sharing, coordination, accuracy, and fast response to information.

Mlo#10 Systems Framework/ESSP Integration

The importance of this project was to help SCVWD streamline their workflow in order to obtain the information they need to complete their work faster, quickers and more accurate. This

project covers the areas of socioeconomic, political, and ecological areas which the SCVWD deal with. First we had to deal with the human component of the system. GIS is widely used through out the District; with more than 130 GIS Users in 50 units using GIS to create and access data. Duplication of data was an issue that had to be resolved for accuracy reasons and productivity reasons, and access reasons. The problem with data created this way is because of the use of the data. GIS data is used for ecological reasons when dealing with delineating habitat areas, or when dealing with parcel information for tax purposes. Data is also used to show the board members and citizens of Santa Clara County. The data is used to convince public of policy decisions, or project, or funding.

The focus of this project was to focus on providing the best available data, and getting rid of duplicate efforts, and create metadata to show how the data was created and it's accuracy. Major components lift out were the governance structure of the GIS system. This task is overwhelming and needs support form senior management, and is far beyond the skills currently held by the staff of the GIS Administration. This task is planned to be tackled as this proposal is being writing. As a start, Metadata is what is currently being focused on.

Key assumptions we are making are that people will come forward and claim their GIS Data.
The system will effect ecological components by keeping track of past, current, and future trends by looking up the information through Metadata. It will also effect human stakeholders by allowing them to accomplish their jobs faster and easier, with accurate and current data, also know where the information came from. As a result, people will be able to make better decisions by using better data.

~~MLO #11 Service Learning ???~~

MLO #4 Application of Economic/Political Knowledge.

MLO #5 Acquisition, Display, and Analysis of Quantitative Data.

9. **Interview David Painter**

The interview with David Painter should be informal and informative. The purpose of this interview is to get information from Mr. Painter that supports the implementation of the Federal Geographic Data Committee Metadata Standard. Some questions may focus on Mr. Painter's view on the importance of Metadata; some of the policy implications going on; new problems that may arise if policy isn't passed; and the future outlook of Metadata.

New CAP

- cover letter (download from net)
- Summary / Abstract
    - 1 paragraph – what you plan to do.

- 1-2 sentences on the general question, problem, or topic CAP will address- the big picture;
- a description of the project;
- the goals of the project;
- the methods that will be used;
- how project is related to the general question, problem, or topic that inspired it (enterprise)

- Introduction / or new / ? Systems Diagram ?
  - Show the interaction among the relevant socioeconomic, political, ecological components CAP focuses on.
- Background and Goals
  - 3-6 pages typed / double spaced
  - General interactive ESSP systems context which CAP is embedded.
  - Identify a General problem (environmental ? Data ?...)
  - Framework in which problem exist;
    - General Socioeconomic, political, ecological/physical components of the system.
    - General interactions within and among the components of the system.
  - Identify/describe the specific ecological/physical and human components of the system.
    - Present scientific info that increases the understanding of the issue; identify any key gaps in the scientific knowledge.
    - Identify relevant policies (existing, proposed federal, state, local, organizational policies-lack of policies); and/or behaviors/attitudes relevant to the issue.
    - Describe interaction between ecological/ physical/human (socioeconomic, political) components of the system, and identify specific areas your work focuses on.
  - Critically evaluate strengths and limitations of your description of the system.
    - Identify major components you left out of description and why you left them out.
    - Identify key assumptions you are making.
  - Evaluate the effects of alternative policies or human actions on the behavior of the system.
    - Explain how it will effect ecological/physical components under different scenarios (i.e. policies).
    - How it will effect human stakeholders under different scenarios. (i.e. policies).
  - Identify the purpose of the CAP project, and link it to the general system you have described.
    - Name the specific hypothesis you plan to test; question you plan to answer; goal you hope to reach.
    - Relate the question/goal of project to the environmental issue.
    - Explain what effects you hope project will have on the environmental issue.
  - Describe General approach you will take to test, answer, or achieve goal.

- Methods
    - 2 pages
    - explain in as much detail exactly what you will be doing for the project.
    - Very detailed.

- Areas of Depth
    - 1 paragraph
    - which areas of depth you want to be graded on.
    - Why do you feel these are important.

- Real world applications, ethics, personal bias.

- Timeline

- Literature Cited