

University of Redlands

InSPIRe @ Redlands

---

MS GIS Program Major Individual Projects

Theses, Dissertations, and Honors Projects

---

12-2018

## Legionnaires' Disease - A Prototype Tool to Identify At-Risk Hospital Rooms

Jonathan T. Quinn

Follow this and additional works at: [https://inspire.redlands.edu/gis\\_gradproj](https://inspire.redlands.edu/gis_gradproj)

 Part of the [Geographic Information Sciences Commons](#)

---

### Recommended Citation

Quinn, J. T. (2018). *Legionnaires' Disease - A Prototype Tool to Identify At-Risk Hospital Rooms* (Master's thesis, University of Redlands). Retrieved from [https://inspire.redlands.edu/gis\\_gradproj/278](https://inspire.redlands.edu/gis_gradproj/278)



This work is licensed under a [Creative Commons Attribution 4.0 License](#).

This material may be protected by copyright law (Title 17 U.S. Code).

This Thesis is brought to you for free and open access by the Theses, Dissertations, and Honors Projects at InSPIRe @ Redlands. It has been accepted for inclusion in MS GIS Program Major Individual Projects by an authorized administrator of InSPIRe @ Redlands. For more information, please contact [inspire@redlands.edu](mailto:inspire@redlands.edu).

University of Redlands

**Legionnaires' Disease – A Prototype Tool To Identify At-Risk Hospital  
Rooms**

A Major Individual Project submitted in partial satisfaction of the requirements  
for the degree of Master of Science in Geographic Information Systems

by

Jonathan Quinn

Mark Kumler, Ph.D., Committee Chair

Ruijin Ma, Ph.D.

December 2018

Legionnaires' Disease – A Prototype Tool to Identify At-Risk Hospital Rooms

Copyright © 2018

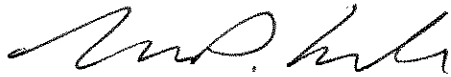
by

Jonathan Quinn

The report of Jonathan Quinn is approved.

A handwritten signature in black ink, appearing to be 'Ruijin Ma', written over a horizontal line.

Ruijin Ma, Ph.D.

A handwritten signature in black ink, appearing to be 'Mark Kumler', written over a horizontal line.

Mark Kumler, Ph.D., Committee Chair

Dec 2018



## **Acknowledgements**

I would first like to thank my parents who consistently give me the motivation to better myself. I would also like to extend thanks to Esri for funding my education, my team members for accommodating my schedule, and the MS GIS faculty and adjunct faculty at the University of Redlands. I also appreciate the support I received from my advisor, Dr. Mark Kumler, Dave Crawford for assisting on some specifics about geometric networks, Prachi Asher at the Veteran Affairs Hospital for the inspiration, and Esri's Facilities team for providing some important resources.

Most importantly, I couldn't have done any of this without the patience and support from my wife, Jill. She helped me stay on track and focused on obtaining a Master's degree.



## **Abstract**

Legionnaires' Disease – A Prototype Tool to Identify At-Risk Hospital Rooms

by

Jonathan Quinn

Nosocomial diseases are diseases that originate within hospital or healthcare facilities. Those types of diseases pose significant risks to patients staying within hospital or healthcare facilities as patients are often elderly or already have compromised immune systems. Often, the diseases are airborne or waterborne and frequently spread through the building's utility systems, such as the heating, ventilation, and air conditioning, (HVAC), or water system. This report describes a proof of concept to identify parts of an HVAC system that may be affected by a contaminate, and which rooms within the building are at risk of spreading the contaminate. The implementation involved converting CAD data to file geodatabase feature classes which are in turn used to build a geometric network. Tracing analysis is then used to identify rooms and HVAC vents likely affected by a contaminate based on an input room. The tool developed for this project can both identify which rooms may be affected by a pollutant but also generate a list of rooms to sample when developing a preventative testing schedule.





# Table of Contents

<b>Chapter 1 – Introduction .....</b>	<b>1</b>
1.1 Client.....	1
1.2 Problem Statement .....	2
1.3 Proposed Solution .....	2
1.3.1 Goals and Objectives .....	3
1.3.2 Scope.....	3
1.3.3 Methods.....	3
1.4 Audience .....	5
1.5 Overview of the Rest of this Report .....	5
<b>Chapter 2 – Background and Literature Review .....</b>	<b>7</b>
2.1 <i>Legionella pneumophila</i> .....	7
2.1.1 Dispersion and Transmission.....	8
2.1.2 Water System Treatment.....	9
2.2 Georeferencing CAD Data.....	10
2.3 Geometric networks .....	11
2.4 Analysis of Disease Movement .....	11
2.5 Summary .....	12
<b>Chapter 3 – Systems Analysis and Design.....</b>	<b>15</b>
3.1 Problem Statement .....	15
3.2 Requirements Analysis .....	15
3.2.1 Functional Requirements .....	17
3.2.2 Non-functional Requirement .....	18

3.3	System Design .....	18
3.4	Project Plan .....	19
3.4.1	Project Phases .....	19
3.5	Summary .....	21
<b>Chapter 4 – Database Design.....</b>		<b>23</b>
4.1	Conceptual Data Model .....	23
4.2	Logical Data Model .....	24
4.3	Data Sources .....	25
4.4	Data Scrubbing and Loading .....	25
4.5	Summary .....	25
<b>Chapter 5 – Implementation.....</b>		<b>27</b>
5.1	CAD Data Conversion .....	27
5.1.1	Geofencing CAD Data.....	27
5.1.2	Isolating Duct Lines .....	30
5.1.3	Converting Polylines to Centerlines .....	31
5.1.4	Creating Sources and Extracting Sinks.....	35
5.2	Creating the Geometric Network .....	37
5.2.1	Validating the Geometric Network.....	38
5.3	Developing the Analysis and Web Products.....	39
5.3.1	Building the Geoprocessing Script .....	40
5.3.2	Developing the Web Products .....	42
5.4	Summary .....	43
<b>Chapter 6 – 44</b>		

<b>Chapter 7 – Results and Analysis.....</b>	<b>Error! Bookmark not defined.</b>
7.1 Tracing Indoor Utility Features .....	<b>Error! Bookmark not defined.</b>
7.2 Converting CAD Data.....	<b>Error! Bookmark not defined.</b>
<b>Chapter 8 – Conclusions and Future Work.....</b>	<b>49</b>
8.1 Conclusion .....	49
8.2 Future Work .....	50
<b>Works Cited .....</b>	<b>53</b>
<b>Appendix A. Geoprocessing Script for Tracing.....</b>	<b>58</b>



## Table of Figures

Table 1 Functional and Non-functional Requirements.....	16
Figure 3-1 System Design.....	18
Figure 3-2 Project Phases.....	19
Figure 4-1 Conceptual Data Model.....	24
Figure 5-1 Default CAD Alignment .....	28
Figure 5-2 CAD Control Points .....	29
Figure 5-3 Reference Data Control Points.....	29
Figure 5-4 Georeferenced CAD Data .....	30
Figure 5-5 Isolating Duct lines .....	31
Figure 5-6 Separating Overlapping Features .....	32
Figure 5-7 Selecting and Deleting Extraneous Features.....	33
Figure 5-8 Deleting Erroneous Features .....	34
Figure 5-9 Polygon to Centerline Output.....	35
Figure 5-10 Select Vent Features.....	36
Figure 5-11 Validating Digitized Direction.....	37
Figure 5-12 Validating Digitized Direction.....	39
Figure 5-13 Script Tool Design .....	40
Figure 5-14 Upstream Trace Analysis .....	41
Figure 5-15 Downstream Trace Analysis Results .....	42
Figure 5-16 Web Application Results.....	43
Figure 6-1 Rooms Identified From Downstream Trace Results.....	46
Figure 6-2 CAD Data Model .....	47

Figure 6-3 Selecting for Vents .....	48
Figure A 1 Trace Features (Lines 1-31) .....	58
Figure A 2 Trace Features (Lines 31-62) .....	59
Figure A 3 Trace Features (Lines 62-93) .....	59

## **List of Tables**

Table 1 Functional and Non-functional Requirements.....	16
---	----





## **List of Acronyms and Definitions**

DWG	CAD Drawing file
HVAC	Heating, Ventilation, and Air Conditioning



# Chapter 1 – Introduction

When a patient is receiving care from a hospital or healthcare facility, their intention is to get better, not worse. Unfortunately, airborne or water borne nosocomial diseases, meaning ones that originate within hospitals, pose significant risks to patients within health care settings. Since it is common that those types of diseases spread via utility systems, such as heating, ventilation, and air conditioning systems or water systems, it is important that a hospital or healthcare facility be able to quickly identify and contain the source of the contamination. It is also important to actively monitor and sample from potential outlets to ensure that any contaminations are quickly abated before they become significant issues.

This paper describes a framework that can be used to identify rooms within a building that may be at risk of spreading contaminants. By isolating utility features and rooms that may be contaminated, a hospital or healthcare facility can quickly identify the source of the contaminate and quickly contain it. Section 1.1 describes the client for this project and Section 1.2 describes the client's problem. Section 1.3 covers the proposed solution. The intended audience is described in Section 1.4 and finally, an overview of the rest of the paper is outlined in Section 1.5.

## 1.1 Client

The client for this project is the Veteran Affairs (VA) Hospital in Loma Linda, California. The hospital is a four-story building that contains around 3,200 rooms. The VA has had reported cases of Legionnaires' disease, caused by *Legionella pneumophila*. Legionnaires' disease is commonly spread through building plumbing systems and

untreated groundwater (Beer, 2015). The presence of Legionnaires' disease is tested by sampling the water at different outlets, such as sink faucets and showers, and treatment is done on the entire water line.

## **1.2 Problem Statement**

According to analysis performed by the Center for Disease Control and Prevention, the most common outbreak settings for Legionnaires' disease include hospitals and long-term care facilities (Garrison, et al., 2016). Recently, patients staying at the VA Hospital have contracted Legionnaires' disease. Since the disease is spread through the water supply, and given that many of the 3,200 rooms at the hospital have water outlets, the hospital would like a systematic approach to testing the water lines. Instead of randomly testing all water lines on a floor, they would like to isolate specific water loops. By identifying the water lines or areas within the hospital that have been affected, the VA can quickly apply isolated mitigation techniques while minimizing mitigation costs. This project aims to provide a framework and proof of concept on identifying water lines and sources of water that may be infected with *Legionella pneumophila* once such a case occurs.

## **1.3 Proposed Solution**

The proposed solution is to isolate water lines using geometric network tracing analysis that will allow the hospital to identify lines that are upstream of rooms where patients have contracted the *Legionella pneumophila* bacteria. Once the analysis is run, the hospital would have identified specific water lines to test rather than needing to test all water lines throughout each floor. This will aid in identifying patients that may be affected, and minimize testing and treatment costs as the hospital will have narrowed

down water lines that are affected. The solution will be developed with sample data, which serves as a proof of concept for the hospital.

### **1.3.1 Goals and Objectives**

By isolating specific water lines that are connected to rooms where patients have contracted Legionnaires' disease, the hospital can focus treatment efforts on those water lines as opposed to treating the whole water system. This effort will save money and time as the hospital can address specific water lines. It will also provide a tool for the hospital if similar situations arise in the future. This project will provide a proof of concept by identifying major steps and tasks required to develop an analysis that can isolate water lines.

### **1.3.2 Scope**

The scope of the project was relatively narrow, involving migrating CAD floor plans and water line drawings to a file geodatabase, building a geometric network of the migrated data, and identifying upstream water lines from hospital rooms where patients have contracted Legionnaires' disease.

### **1.3.3 Methods**

To complete the project, there were three primary phases. The first phase involved finding CAD floor plans and water lines data and migrating that data to a file geodatabase. Each floor plan had to be aligned to the correct location through imagery and the water lines aligned to the floor plans. This data would then be incorporated into a geometric network for downstream and upstream tracing. The second phase of the project was to develop the analysis to properly identify upstream water lines from instances of

Legionnaires' disease, as well as identify downstream water lines from affected sources. The final phase was to publish a geoprocessing service and build a web application that used this service so non-GIS users can run the analysis.

The first phase of the project required finding suitable CAD data to use as a prototype as data from the VA was not readily available. Esri's facilities team was contacted and they were able to provide the floor plan data, but the team did not have water line data. However, Esri did have heating, ventilation, and air-conditioning (HVAC) data instead. The HVAC data was used in this project to prove the concept. Once CAD data was acquired, the data was migrated to file geodatabase feature classes. This involved first georeferencing the CAD data against reference data. Once the data aligned correctly, the data was converted from CAD to file geodatabase feature classes and the NAD 1983 State Plane California VI FIPS 0406 Feet projection assigned to the converted data. The logical data model was constructed in the first phase, removing extraneous fields from the CAD drawings as well as adding and updating field names and values for clarity.

Once the data was prepared and converted, the second phase involved creating a geometric network from the data and developing the analysis through Python with the arcpy module. Sources and sinks were defined for the data and the geometric network was built. A script tool was developed to trace upstream ducts based on input room numbers. The source AC unit returned by the ducts was then used to trace downstream to identify the rooms that the AC unit served.

The final phase was to publish a geoprocessing service and develop a web application that included a geoprocessing widget that called the service. This ensured that any non-GIS users without knowledge of ArcMap or ArcGIS Pro could run the analysis.

## **1.4 Audience**

The audience for this project are primarily the staff at the Loma Linda VA Hospital, but the project can apply to a larger audience of healthcare officials interested in ways to analyze utilities within their setting. This type of analysis can be applied to other water or airborne diseases other than Legionnaires' disease. The analysis can aid in identifying outlets or areas to test, thus making testing more isolated and effective. The project does involve moderately advanced workflows in GIS as well as knowledge of how water lines are laid out within a building, so additional audience members may include users knowledgeable in architecture as well as GIS administrators.

## **1.5 Overview of the Rest of this Report**

The rest of the report is split into several large categories. Chapter two describes background information and literature review on Legionnaires' disease along with information on the *Legionella pneumophila* bacteria. It will provide a discussion on topics important to this project, such as dispersion methods and settings it is more likely to thrive in. Chapter two will also describe georeferencing CAD data as well as background on geometric networks. Chapter three will go over system analysis and design to describe the problem addressed by this report, the proposed solution and major tasks to complete the project. Chapter four will go over the database design. It will cover both the conceptual model of how the relationships between spatial objects are defined as well as the logical model of physical properties of the database. Chapter five covers the



implementation of the project plan. It will cover methodologies and tasks completed, such as data conversion and analyses conducted. Chapter 6 describes the results of the project and reflects on how the project progressed. Chapter 7 summarizes the project and describes future work to expand on what was accomplished.

## **Chapter 2 – Background and Literature Review**

Between 2000 and 2014, reported cases of Legionnaires' disease have increased from 0.42 per 100,000 persons to 1.62 per 100,000 persons (Garrison, et al., 2016). Garrison, et al. (2016) further determined that between 2000 and 2014, health care associated outbreaks accounted for over half, (57%) of the 27 investigated outbreaks and accounted for 85% of the deaths attributed to Legionnaires' disease. Considering the elderly and individuals with compromised immunodeficiency syndromes are more likely to contract Legionnaires' disease, it is understandable that patients in health care facilities would be at increased risk of contracting the disease (Marston, Lipman, & Breiman, 1994). The bacteria behind Legionnaires' disease, *Legionella pneumophila*, typically spreads through "water cooling devices, large buildings with wet cooling systems used for air conditioning, hot and cold-water systems used in hospital" (Heuner & Swanson, 2008, p. 37). The combination of a large density of restrooms and showers within hospital rooms, and patients having compromised immune systems, causes Legionnaires' disease to be particularly dangerous.

### **2.1 *Legionella pneumophila***

Legionnaires' disease is a potentially fatal type of pneumonia caused by *Legionella pneumophila* and has gained attention after an outbreak of pneumonia among participants who attended the 1976 American Legion Convention in Philadelphia, (Bartram, Chartier, Lee, Pond, & Surman-Lee, 2007). 182 attendees contracted pneumonia like symptoms that resulted in 29 deaths (Heuner & Swanson, 2008). Due to the setting at a convention for the American Legion, the disease was named Legionnaires' disease. Since then, retrospective analysis of pneumonia outbreaks have been attributed to Legionnaires'

disease, such as considerable outbreaks in 1974, (Terranova, Cohen, & Fraser, 1978), and 1947 (McDade, Brenner, & Bozeman, 1979). Hospitals have developed guidelines for prevention and treatment as Legionnaires' disease is fatal for about one in 10 persons who become infected (Garrison, et al., 2016).

### **2.1.1 Dispersion and Transmission**

Legionnaires' disease found in hospitals is primarily dispersed through the water system and promoted by the “temperature, configuration and age of the hot water tank, physicochemical constituents of the water, plumbing materials, and commensal microflora” (Lin, Stout, Yu, & Vidic, 1998, p. 147). Hot water tank temperatures between 45°C and 50°C provide *Legionella pneumophila* with optimal growing conditions (Lin, Stout, Yu, & Vidic, 1998). The bacteria was typically absent in testing environments when the water temperature was 60°C (Rogers, Dowsett, Dennis, Lee, & Keevil, 1994). It is likely that a hospital's interest in energy conservation by decreasing the sustained water temperature within its water systems may increase the risk of *Legionella pneumophila* colonies being present within the water supply. Many hospitals also maintain “hot water temperatures at levels compatible with patient comfort” (Mandel, Sprauer, Sniadack, & Ostroff, 1993, p. 642), which may not be high enough to neutralize *Legionella pneumophila*.

Legionnaires' disease is not transmitted person-to-person, but rather through “inhalation of contaminated aerosols or aspiration of contaminated water” (Beauté, et al., 2012, p. 2). According to Bartram, Chartier, Lee, Pond and Surman-Lee (2007), once the water system has been contaminated with *Legionella pneumophila*, “warm-water plumbing systems, air-conditioners,” (p. 30) and equipment such as “cooling towers,

building water systems, respiratory therapy equipment and hot tubs” (p. 30) can spread aerosols contaminated with *Legionella pneumophila*. When the aerosols evaporate, they leave particles of less than 5 µm in diameter that can be inhaled, causing legionellosis (Fitzgeorge, Baskerville, Broster, Hambleton, & Dennis, 1983).

### **2.1.2 Water System Treatment**

There are various methods to treat and prevent colonization of *Legionella pneumophila* within water systems. Some of the more common approaches are copper-silver ionization, superheat-and-flush, and hyperchlorination. However, there is typically an inconvenience associated with the treatment methods, such as risk to patients or water lines, or costs associated with the treatment.

Copper and silver ions have been effectively used to kill *Legionella pneumophila* in more than 30 hospitals around the United States (Lin, Vidic, Stout, & Yu, 1998).

Treating water systems with copper and silver ions has also proven to keep concentrations of *Legionella pneumophila* low for over 3 months after treatment was stopped (Liu, et al., 1994).

Using the superheat-and-flush method, temperatures in hot-water storage tanks is increased to 70° C and outlets flushed for 30 minutes (Stout, Lin, Goetz, & Muder, 1998; Lin, Vidic, Stout, & Yu, 1998). The superheat-and-flush method was effective at lowering concentrations of *Legionella pneumophila*; however, it has a tendency to recolonize if the superheat-and-flush is the only method of treatment (Lin, Vidic, Stout, & Yu, 1998).

Hyperchlorination of a water system adds chlorine to the water supply in two ways: continuously and through a shock treatment. Shock treatment (raising the ppm of

free chlorine within the water supply to 20-50), is typically followed by continuous hyperchlorination to keep the free chlorine levels at 0.5-1.0 mg/L (Orsi, et al., 2014).

Hyperchlorination of water systems has met

increasing disfavor because of inadequate penetration of the agent into biofilms in piping, persistence of *Legionella pneumophila* organisms in hyperchlorinated systems, corrosion of the water distribution system leading to pinhole leaks over time, and the introduction of carcinogens into the drinking water (Lin, Stout, & Yu, 2011, p. 170).

## **2.2 Georeferencing CAD Data**

When converting or migrating data into a GIS, the data may come from a variety of different sources. Some of these sources may not support coordinate systems or the data are drawn without respect to a geographic coordinate system. To ensure that converted data align with existing data, the CAD data must be georeferenced first. This aligns the data to reference data, so it can be converted to GIS formats. CAD data often are not digitized according to their geographic locations, but rather in a local coordinate system, or “locations relative to an arbitrary geometric origin, (0,0,0)” (Maher, 2018; ESRI, 2018c). Therefore, when the CAD data are brought into a GIS, such as ArcMap or ArcGIS Pro, they do not line up in their expected location. Georeferencing, (the act of adjusting the location of CAD data or imagery), can help aligning CAD data. Since CAD data are based on engineering and architectural plans and the reference data is likely geometrically accurate, (for example a building footprint or imagery), two control points are typically enough to georeferenced the data. Georeferencing CAD data uses the similarity transformation (ESRI, 2018c). Similarity transformation can “scale, rotate, and

translate data” without skewing by “maintaining the relative shape of the features” (ESRI, 2018a) which means once the data is georeferenced, the shapes of features will match the original CAD drawings.

### **2.3 Geometric networks**

Geometric networks provide the ability to model “connectivity between edges, (such as lines), and junctions where they connect” and are used to perform quick tracing tasks on networks (Design Patterns - Linear Networks, 2004, p. 24). Geometric networks model “real world infrastructure behaviors” by defining topological relationships between connected features and simulating “directed flow system applications” (Services, 2018, pp. 1-6). Using a geometric network, a water utility company can determine where to shut off valves when a pipe bursts by tracing network elements upstream or downstream from the burst pipe (ESRI, 2018b). Environmental monitoring stations can determine the source of pollution within a stream by calculating the shortest path upstream from two points (ESRI, 2018b). The ability to trace upstream features within a geometric network is going to be particularly relevant to this project. In order to trace features within a geometric network, flow direction must be established relative to the source (where the flow originates) to the sinks (where the flow ends). Due to data access issues, water line data was not available from the VA. Instead HVAC data for a building at Esri was used. Given that both are utility systems and a geometric network is designed to model those types of systems, an HVAC system was a suitable alternative.

### **2.4 Analysis of Disease Movement**

Using GIS and spatial analysis has been increasingly important ever since John Snow identified contaminated water as the source for cholera outbreaks in Soho, London,

between 1853 and 1854 that killed around 600 people (Paneth & Paul, 2013). Snow had mapped the locations of patients that had contracted cholera and determined they centered around a specific pump, the Broad Street Pump (Paneth & Paul, 2013). Once the pump handle was removed, new cholera cases subsided. Since that time, there have been significant advances in analytical capabilities, both spatial and temporal, that have aided in disease and pollutant identification, prevention and treatment. Shiode, Shiode, Rod-Thatcher, Rana, & Vinten-Johansen (2015) expanded on Snow's original hypothesis by including population density and spatiotemporal analysis of the cholera deaths. Results from kernel density analysis and Network-based Scan Statistics, which analyzed clustered events confined by street paths, supported the original findings by Snow (Shiode, Shiode, Rod-Thatcher, Rana, & Vinten-Johansen, 2015).

Applications of GIS in disease or pollutant analysis range from large scale (countries and regions), to medium scale (cities and city areas), to small scale analysis. A study by Liu, Hung, Kuo, & Liang (2015) using kriging methods to analyze pollution within an office building. Kriging is an interpolation method "in which the surrounding measured values are weighted to derive a predicted value for an unmeasured location" (Liu, Hung, Kuo, & Liang, 2015, p. 23). They determined that there were locations within the office building that posed threats to the respiratory system and offered recommendations to mitigate those risks.

## **2.5 Summary**

Ever since the first significantly researched outbreak in 1976, Legionnaires' disease, caused by *Legionella pneumophila*, has retrospectively been associated with a number of pneumonia related outbreaks in communities and health-care facilities. To this day,

Legionnaires' disease continues to be a serious issue within hospitals. The bacteria are difficult to eradicate completely and can recolonize after treatment. Additionally, there are inconveniences associated with most of the treatment methods. The cost of the treatment, detrimental side effects on the systems being treated, and risks to patients who rely on the water system all strengthen the case for a more isolated treatment regimen. It is imperative that hospitals define a continuous treatment schedule to prevent colonization within its water supply.

The high density of water outlets (showers, faucets), and the likelihood that patients have compromised immune systems or are older, mean hospitals are at additional risk from Legionnaires' disease. Marrie et al. (1995) have found that it is important to culture all sources of potable water that a patient who has contracted Legionnaires' disease has been in contact with. To determine all sources of potable water, a healthcare facility must be able to trace where the water is coming from and the best way to do that is through a geometric network. By identifying the room where a patient was staying that has contracted Legionnaires' disease, geometric network analyses can identify the source of that outlet. It can further determine other outlets connected to that specific source. The methods to test for Legionnaires' disease are well documented in other reports (Beauté, et al., 2012), so this report will cover a proof of concept in analyzing a utility system in a building to identify connected features within that system. The workflow to trace features can be applied to any type of utility system that has determinate flow, such as a water system.

Using GIS to analyze disease and pollution is critical to identifying and mitigating issues related to those phenomena. From John Snow's early theories in the 19<sup>th</sup> century to



current applications of GIS (including a revisit to John Snow's analysis), GIS has played an important role in shaping how disease and pollution are managed. The application of GIS in the context of this project further shows the variety of situations it can be used.

## **Chapter 3 – Systems Analysis and Design**

In order to isolate which water lines are affected by *Legionella pneumophila*, it is necessary to develop a data model that provides functionality to trace upstream and downstream lines from certain points. This chapter will provide an overview of the problem statement, list functional and non-functional requirements to complete the project and provide an overview of the system design. It will also cover major project phases along with the tasks associated with each phase.

### **3.1 Problem Statement**

The Veteran Affairs Hospital in Loma Linda has recently had cases of Legionnaires' disease within its patients. The hospital would like a method to identify rooms and water outlets within water loops most likely affected by the disease. Since each floor can have multiple water loops from different source water tanks, the hospital can develop more efficient testing methods if they focus testing within specific water loops rather than testing outlets throughout a whole floor. They can then more easily isolate water tanks that may have *Legionella pneumophila* colonies. This project will provide a framework and proof of concept for identifying rooms and outlets within a utility network in a building.

### **3.2 Requirements Analysis**

The main functional requirements of this project are floor plan and HVAC data, a geometric network built from the data, a script tool to run tracing analyses on the geometric network, and a web application that will run the geoprocessing service created

from the script tool. The functional and non-functional requirements are described in Table 1.

**Table 1 Functional and Non-functional Requirements**

<b>Requirement</b>	<b>Requirement Type</b>	<b>Description</b>
Floor plan and HVAC data	Functional	Floor plan and HVAC CAD data converted to file geodatabase format to be used in a geometric network
Geometric network built on converted CAD data	Functional	The geometric network built from the water line and floor plan data to run traces
Trace upstream and downstream	Functional	A script tool that accepts input rooms to determine connected water lines and rooms
Input rooms parameter	Functional	In order to trace from a particular room, the user needs to input the room number (optional parameter)
Input point feature	Functional	A user can place a point using a feature set to run the trace from a particular room
Web application	Functional	A web application will call on a geoprocessing service published from the script tool
Users with no GIS experience can run analyses	Non-functional	The user interface and experience is intuitive enough so users with no GIS experience understand how to run the tool

### 3.2.1 Functional Requirements

The first functional requirement is floor plan and HVAC ducts data from a building within Esri. This data will be used in-place of the floor plan and plumbing data from the VA Hospital. Floor plan and HVAC ducts data were used to build the geometric network and was necessary to run the tracing analyses. These data would be converted to file geodatabase feature classes to be used within a geometric network.

The second functional requirement was the geometric network built from the converted CAD data. A geometric network provides the ability to trace either upstream (against flow towards sources), or downstream (with flow towards sinks), to identify which duct line features are connected to input rooms and their source. For the hospital, this would identify water lines that should be tested for the *Legionella pneumophila* bacteria and which water tanks may have been colonized.

The third functional requirement was a script tool that was written to run the analysis and return results from the trace. It accepted room numbers or a digitized point as input to act as the sinks, and then ran upstream tracing to determine which ducts and AC units were connected to the room numbers. From there, additional downstream tracing could be run to determine which other ducts and vents were connected to the affected AC unit. A downstream trace would identify the rooms that may need to be tested.

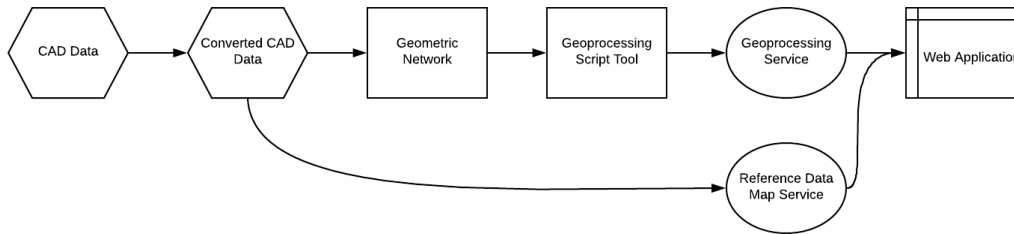
The final functional requirement was a web application to run the analysis. The web application would be accessed from a browser and contain a geoprocessing widget that runs a geoprocessing service created from the script tool.

### 3.2.2 Non-functional Requirement

The one functional requirement for this project was that the web application deliverable should be intuitive and user friendly enough that non-GIS users are able to access it and run the analysis. It is important that the method of running the analysis did not require special software on the users' end, (such as a desktop client like ArcMap and ArcGIS Pro). The web application is simply accessed through a browser.

### 3.3 System Design

The system design is depicted in Figure 3-1.

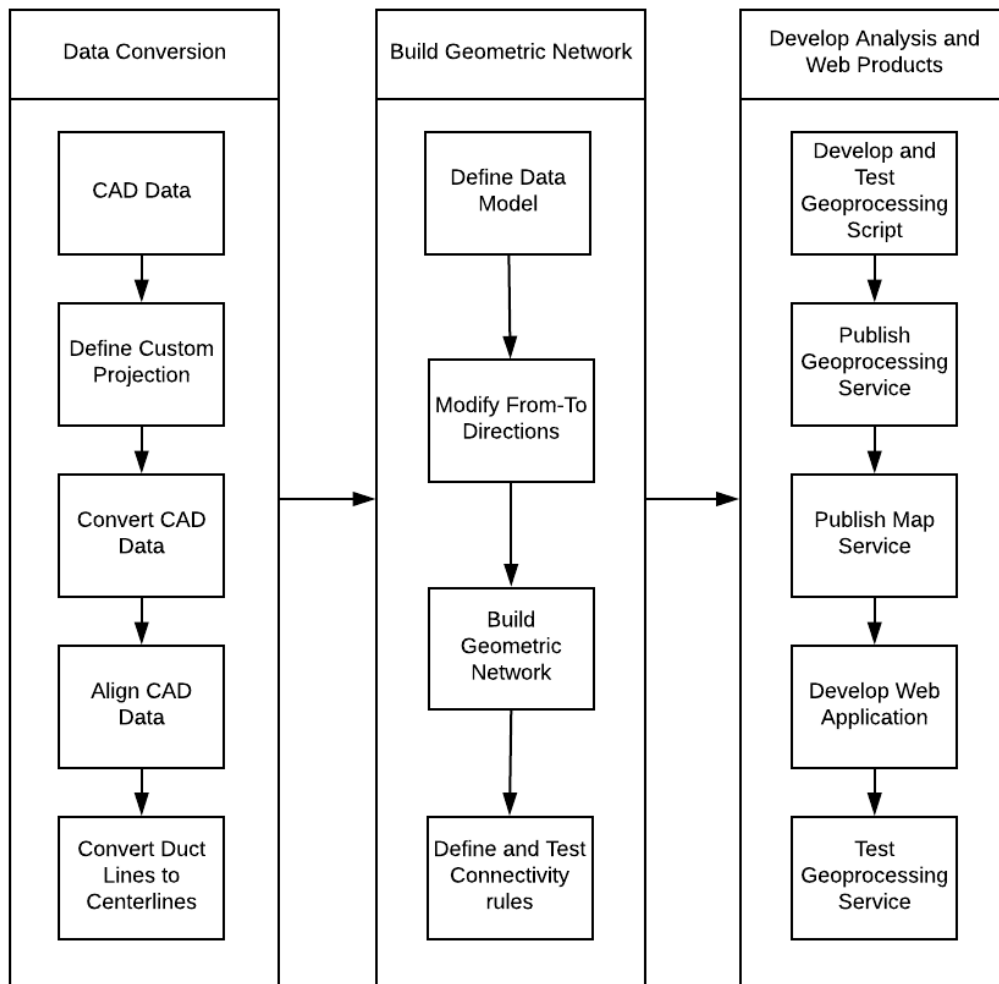


**Figure 3-1 System Design**

There are 7 major components in this project. The first component was the CAD data, which was acquired from Esri. The data depicted the HVAC system within a section of a building on the Esri campus. The next component was the converted CAD data, which was spatially rectified and aligned with existing data. The third component was a geometric network which was built from the converted data. The fourth component was a script tool that ran analyses using the geometric network. The fifth and sixth components were a geoprocessing service published from the script tool and a map service published from the converted CAD data which was to be used as a reference within the final component, a web application. The web application was built through the Web Application Builder.

### 3.4 Project Plan

There were three major project phases, each with individual tasks required to accomplish the major phase.



**Figure 3-2 Project Phases**

#### 3.4.1 Project Phases

The first major phase within the project plan was to convert the CAD HVAC data into file geodatabase feature classes. This involved georeferencing the CAD data as it initially displayed in the middle of the Pacific Ocean. Once the CAD data was georeferenced and

more closely aligned with the building footprint, the data was converted to file geodatabase feature classes. Finally, the ducts polylines were converted into single centerline line features.

The second major phase was developing the geometric network. The first task was to ensure that the attributes and schema of the feature classes were clear, concise and understandable. Once that was done, the from-to directions of all water lines had to be evaluated to ensure proper flow direction. Next, the data was used to build the geometric network and connectivity rules were defined and assessed which involved manually running trace analyses. Once the geometric network was verified, the script tool was developed to accept input room numbers or input points to trace upstream to find which AC units the rooms were connected to. From there, it would optionally also trace downstream to see which rooms the AC unit provided air to. This would simulate the analyses the hospital required; input rooms would be used to trace upstream to hot water tanks and then a downstream trace could be run to see which rooms the hot water tank provided water to.

The final major phase was to develop the web components. This allowed the analyses to be carried out through a user-friendly web application which did not require desktop clients, other than a browser. The first tasks were to publish the geoprocessing service and the map service. The map service was used to provide reference of where the analyses were being run within the building and provide spatial context when the affected water line features were returned. The final task was to develop the web application using the Web Application Builder.

### **3.5 Summary**

Developing a clear project plan with defined phases provided a clear path to a successful project. The project plan helped to prepare the data, workflow and processes for the next steps. Identifying major phases helped to get a big picture of the state of the project, while the individual tasks helped to get a more granular look at how the project was progressing. The plan also helped to ensure that the project methods could be applied to different type of data. Since the analyses involved tracing utility networks, it could be applied to HVAC systems or plumbing within a building.



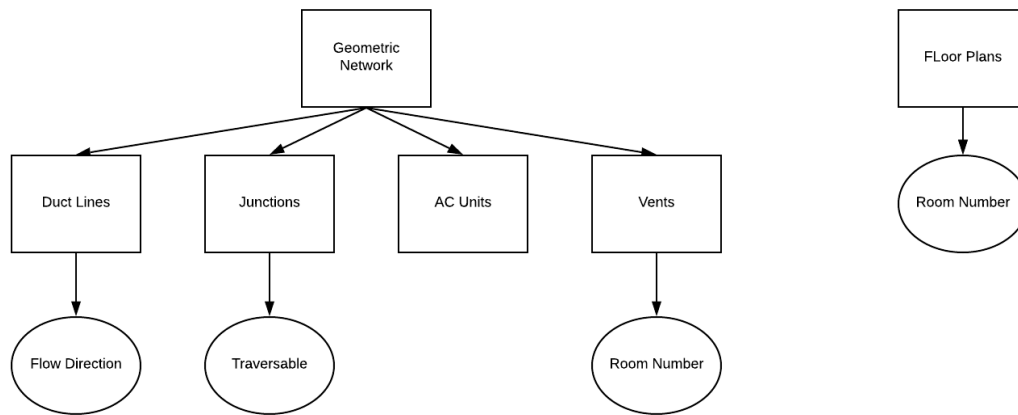


## **Chapter 4 – Database Design**

The database design was an important part of this project as the original data format was CAD, which often stores data differently than a GIS. It was important to convert the CAD data to file geodatabase feature classes. This would allow for advanced functionality, such as a geometric network. The first step to designing a database is defining the conceptual data model. The conceptual data model, explained in Section 4.1 of this report, constructs the different entities within the data model and establishes their relationships. The second step is to design the logical data model, which focuses on the structure of the data itself. This is described in Section 4.2. Section 4.3 outlines the data sources used in this project. Section 4.4 covers how the data was modified to fit the data model and requirements of the project while the last section, 4.5, provides a summary of this chapter.

### **4.1 Conceptual Data Model**

The conceptual data model “captures all of the geodatabase properties at a glance” (Arctur & Zeiler, 2004, p. 12). It provides an overview of the data model at a higher level than the technical implementation of the geodatabase, such as the type of database or schema. In the context of a GIS, the conceptual model describes how spatially related data can be organized (Arctur & Zeiler, 2004). The conceptual model for this project is described in Figure 4.1.



**Figure 4-1 Conceptual Data Model**

## 4.2 Logical Data Model

The logical model for this project consists of the floor plan and converted HVAC data stored in a geometric network. The data was stored in a file geodatabase within a feature dataset per the requirements of a geometric network. The data could be stored in any type of relational geodatabase, such as SQL Server, Oracle, or Postgres.

The AC ducts, AC units, and vents were the data included in the geometric network. The ducts required specific digitized direction to mimic flow direction away from the AC units. They did not have any attribute information prior to adding them to the geometric network. The AC units and vents were two point feature classes used within the geometric network as well. The AC units acted as the source and the vents acted as the sinks for the trace functionality. The vents required an attribute for the room number, which was used as the starting point of the analysis. The room number field name was RoomNumber and was a text field of 25 characters. Once the data was added to a geometric network, a junctions feature class was created. Junctions within a geometric network are used to transfer flow from one edge, such as an AC duct, to another edge.

The floor plan data contained polygons depicting the rooms within the M building on Esri's main campus. The floor plan contained a single text field called RoomNumber that stored the room number. The text fields could store up to 25 characters.

### **4.3 Data Sources**

The floor plan and CAD HVAC data for Building M was provided by Esri to be used as a proof of concept. The data did not have metadata associated with it and was created at least 5 years ago.

### **4.4 Data Scrubbing and Loading**

Two steps were required before the data could be used in a geometric network: georeferencing the CAD data and then defining a projection for the converted data.

Since CAD data is created in a different type of coordinate system than what a GIS would use, the CAD data first had to be georeferenced against building footprint reference data. This brought the CAD data closer to the real-world location. Once it was aligned, it was exported into file geodatabase feature classes. The NAD 1983 State Plane California VI FIPS 0406 Feet projection was used for the feature classes.

The specific implementation of data conversion and loading will be discussed in detail in Chapter 5.

### **4.5 Summary**

This project demonstrates how CAD data can be incorporated into advanced GIS workflows. Special attention needs to be paid to data conversion and attributes of features, as the data model often used in CAD may not work well within a GIS. A common issue that many GIS projects face when working with CAD data is alignment

issues. It is necessary to understand the fundamentals of georeferencing CAD data in order to align it.

## **Chapter 5 – Implementation**

The implementation involved three major phases, each with varying amounts of tasks: converting CAD data to feature classes, building the geometric network, and developing the analysis and web products. The phases were outlined in Section 3.4.1 and depicted in Figure 3-1.

### **5.1 CAD Data Conversion**

The first phase was to convert the CAD data to feature classes. This involved three steps: georeferencing the CAD data to the building footprint, converting the data to file geodatabase feature classes, and then defining the NAD 1983 State Plane California VI FIPS 0406 Feet projection to the converted data. The duct lines were dual polylines and had to be collapsed to a single centerline once the data aligned correctly.

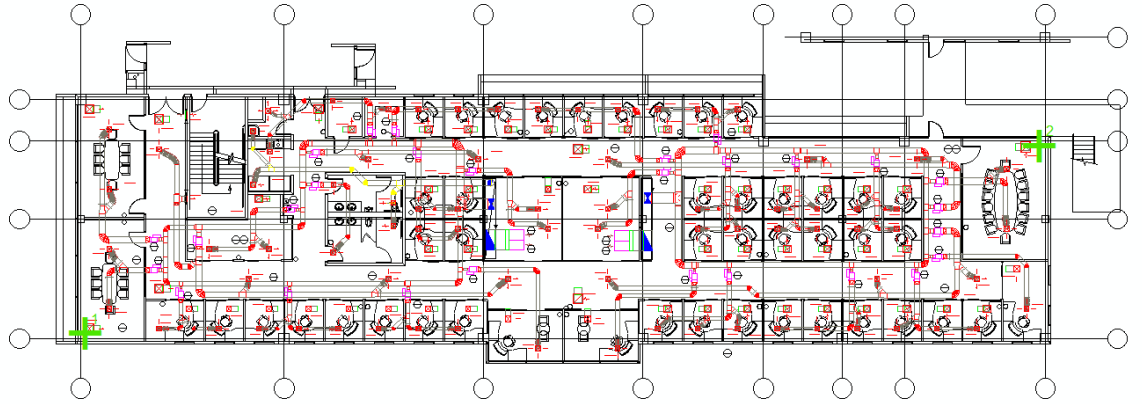
#### **5.1.1 Georeferencing CAD Data**

When CAD drawings are created, they are drawn in a local coordinate system. The coordinate system is not typically a real-world coordinate system and therefore, the features do not line up to real-world features. The data often lines up at 0,0 when displayed over a projected coordinate system, such as Web Mercator Auxiliary Sphere, which can be seen in Figure 5-1.

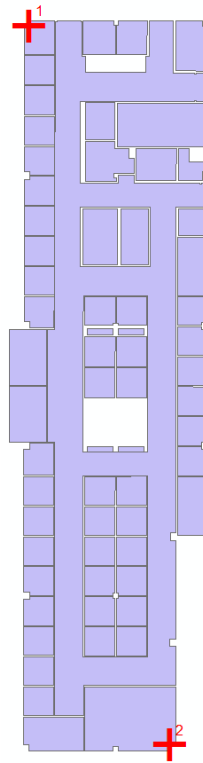


**Figure 5-1 Default CAD Alignment**

Lining up the CAD data requires georeferencing the CAD data using known locations, such as imagery or similar data, as control points. To align the CAD data for this project, the building footprint was used as reference data. Two corners of the CAD data were matched up with corresponding corners of the building footprint (green crosses in the bottom left and top right of the CAD drawing in Figure 5-2 and red crosses in the top left and bottom right of the reference data in Figure 5-3).



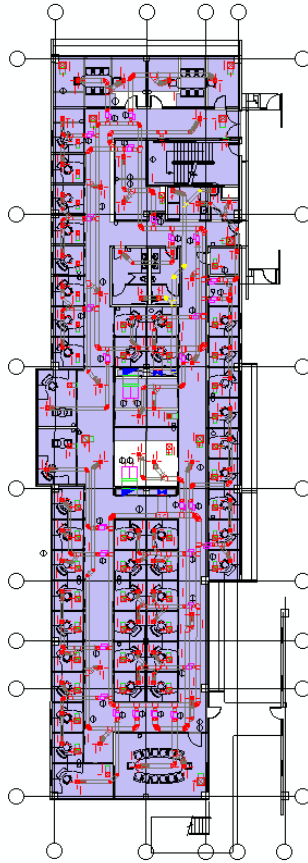
**Figure 5-2 CAD Control Points**



**Figure 5-3 Reference Data Control Points**

Once the data is georeferenced, it can be exported from the CAD drawing into a file geodatabase feature class and the converted data assigned a projection. For this project, the NAD 1983 State Plane California VI FIPS 0406 Feet projection was used as it was the projection appropriate for the spatial extent of the data.





**Figure 5-4 Georeferenced CAD Data**

### **5.1.2 Isolating Duct Lines**

In ArcMap or ArcGIS Pro, CAD drawing files show up as five separate layers under a main group layer: annotation, points, polylines, polygons, and a multipatch layer. The polylines contained the duct data relevant to this project. All polyline features have attribute information based on the CAD drawings, and the attributes specific to the duct lines were extracted using a definition query (Figure 5-5).

Definition Query:

```
("Entity" = '2DPolyline' AND "Layer" = 'MN-SUP-DUCT') or ("Entity" = 'LWPolyline' AND "Layer" = 'MN-SUP-DUCT-HIDE-FLEX') or "Layer" = 'M-EQUIP' OR ( "Entity" IN ('LWPolyline','Line') AND "Layer" = 'MN-SUP-DUCT-HIDE') or "Layer" = '0' and RefName = 'FX4B-60'
```

Query Builder...

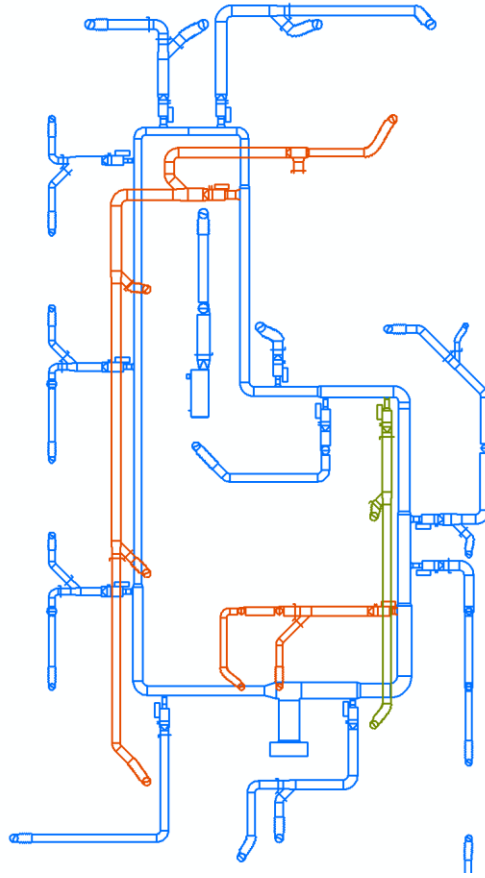
**Figure 5-5 Isolating Duct lines**

CAD drawings are created with a legend and other metadata. When bringing the CAD data into ArcMap, the legend and metadata will appear as though they are features. The definition query ensured that only duct features were selected and exported to a file geodatabase feature class.

### 5.1.3 Converting Polylines to Centerlines

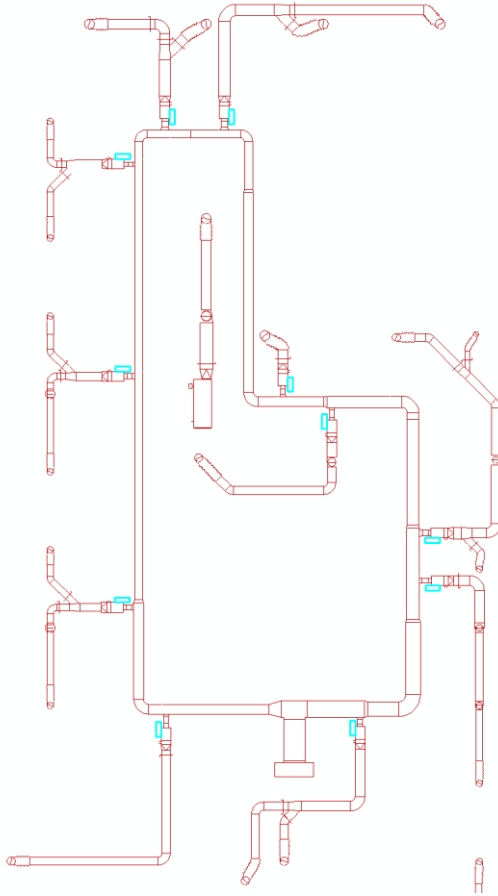
In the original CAD data, the duct polylines were drawn as outlines of polygons. It was important to create centerlines from the polyline features. This involved converting the polylines to polygons, dissolving the polygons to create a single feature, and then converting the polygons to centerlines.

After inspecting the data, the first step was to identify overlapping features. Part of the conversion process was to dissolve polygon features, so it was important that duct geometries remained consistent and were not merged into other duct lines due to overlapping features. Depicted in Figure 5-6, three separate groups of features were identified, and selections made to split the features into separate layers for processing.



**Figure 5-6 Separating Overlapping Features**

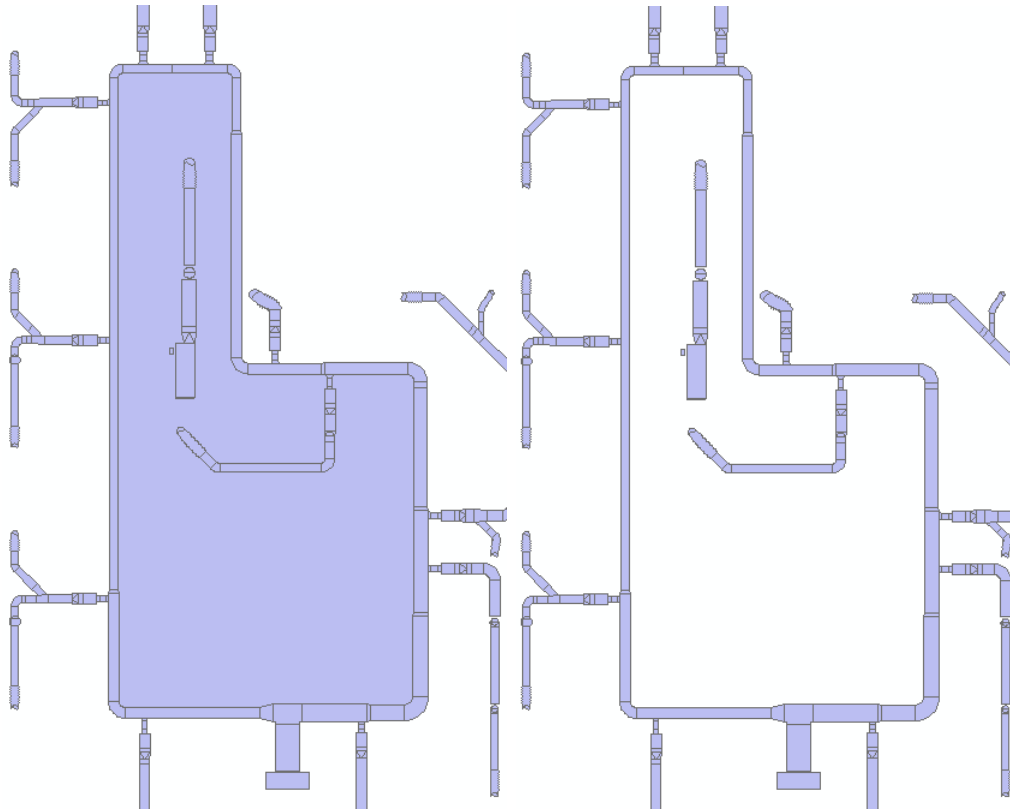
Extraneous features were identified and deleted. CAD drawings are drawn to precise specifications which means feature lengths are determinate. A feature that is a part of a piece of equipment present in multiple places throughout the duct lines will have the same length and geometries wherever it is drawn. If the feature is not needed, it can be selected and removed from the converted data. First, the features in all three layers were converted to single part features using the Multipart to Singlepart tool. Next, a selection was applied to select the extraneous features, given that the shape and geometries of the features were determinate. This selection is represented in Figure 5-8 as light blue features:



**Figure 5-7 Selecting and Deleting Extraneous Features**

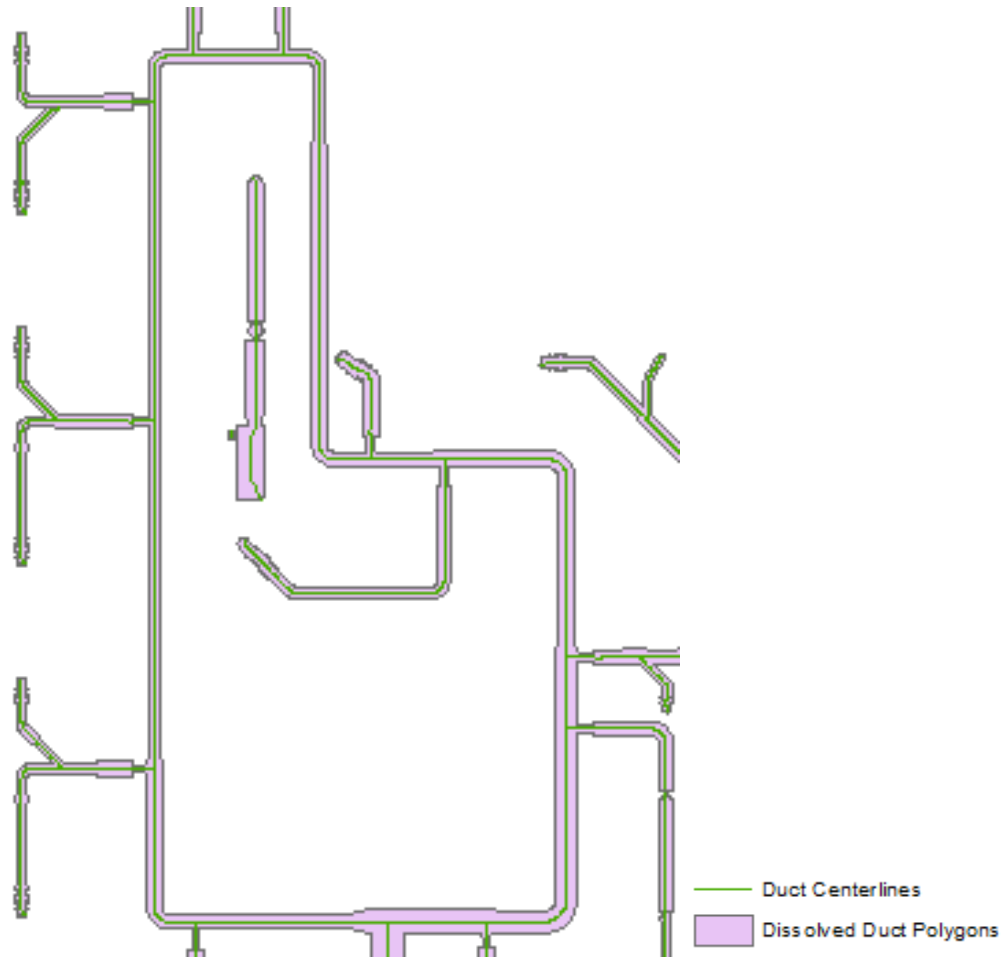
The features were deleted as they did not add any value and could potentially affect the geometries of the duct lines.

After extraneous features were deleted, the data was converted into polygons using the Feature to Polygon tool. Features that were clearly erroneous were deleted, such as a line feature that did not close completely, creating a large polygon which is shown in Figure 5-8.



**Figure 5-8 Deleting Erroneous Features**

With the unneeded and erroneous features deleted, the data was dissolved. This was an important step as the Polygon to Centerline tool available in ArcGIS Pro worked best with dissolved polygons. The output was a single line feature depicting the centerline of all polygons. Figure 5-9 shows the green centerlines displayed over the dissolved polygons.



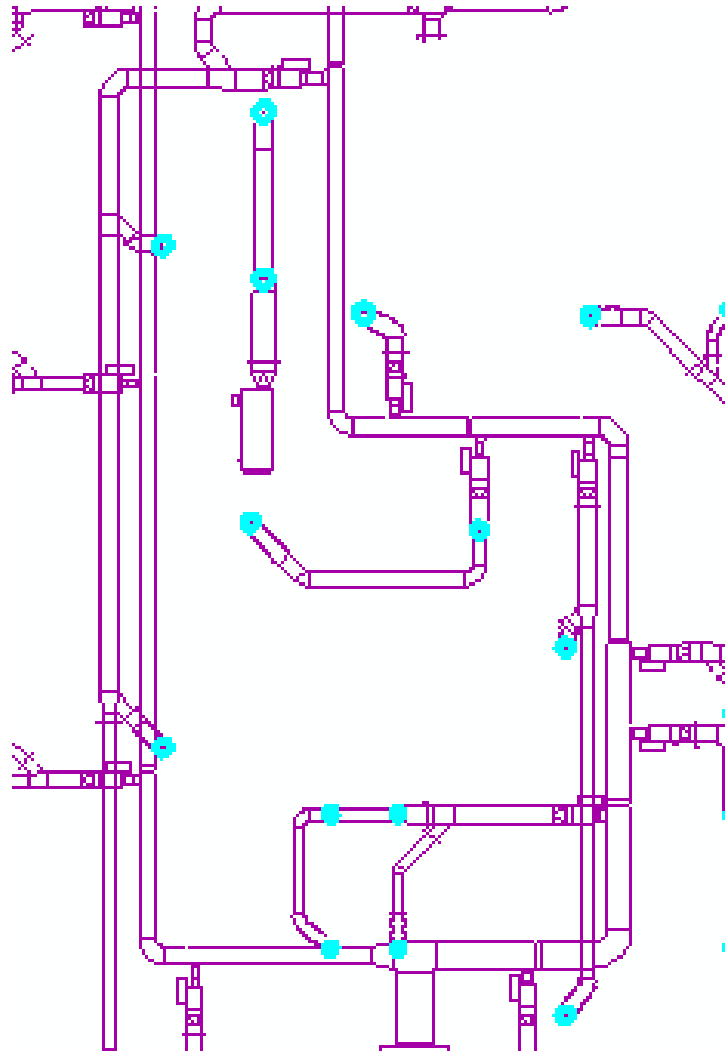
**Figure 5-9 Polygon to Centerline Output**

Once the three separate layers were converted to centerlines, they were merged to create a single feature class using the Merge tool.

#### **5.1.4 Creating Sources and Extracting Sinks**

The AC unit locations were digitized manually based on the CAD drawings to act as the sources. The vents were extracted from the CAD data to act as sinks. The vents were depicted as line features within the original CAD data, so it was first necessary to define a SQL statement that selected those lines. Since CAD drawings are drawn precisely, there was an expected range for the length of the features depicting vents. An attribute query

along with a geometry query allowed the SQL statement to only select the vents' features which can be seen as cyan features in Figure 5-10.



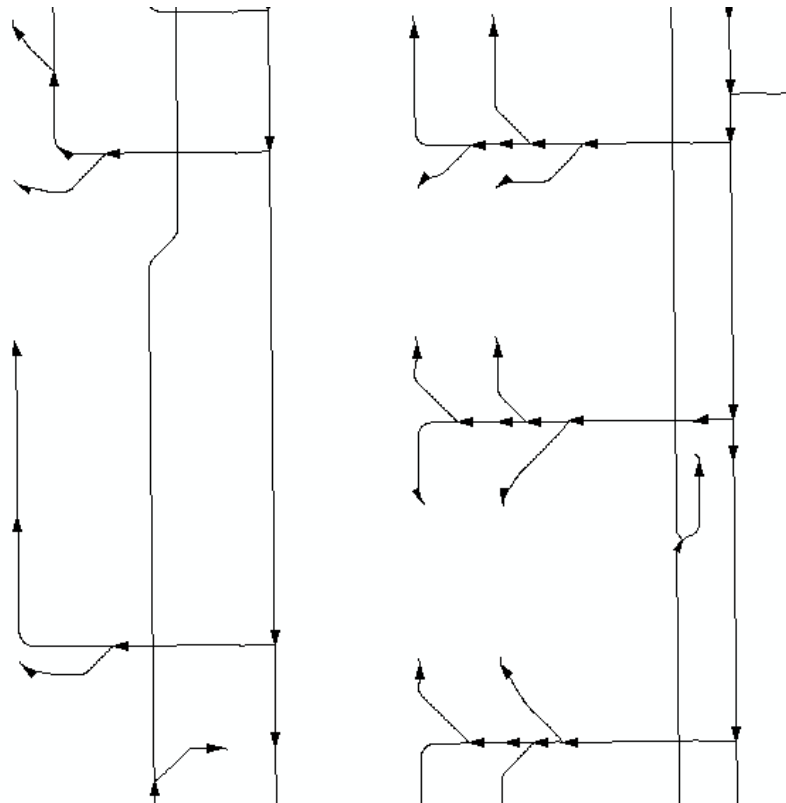
**Figure 5-10 Select Vent Features**

Once the vent features were selected, the Minimum Bounding Geometry tool was run to convert the features into Polygons. Next, the Feature to Point tool converted the polygons to points. The resulting points were the vents within the HVAC system. The

vents were then snapped to the ducts' centerlines using the Snap tool within the Editing Tools toolset to make sure that the vents can be used as input flags for tracing.

## 5.2 Creating the Geometric Network

Correct flow direction is required to do tracing analysis. Symbolizing line features by their digitized direction is a good technique to ensure that the direction of flow is correct which can be seen in Figure 5-11. After the centerline ducts were symbolized based on their digitized direction, there were a few that had to be modified manually. This was accomplished by entering an edit session, selecting all lines that had to be updated, and using the right-click context menu to flip lines.



**Figure 5-11 Validating Digitized Direction**

Once the ducts and vents were converted, their geometries were verified, and the AC units were digitized, the geometric network was built. The default values and



parameters for a geometric network were used and the ducts, vents, and AC units feature classes were included within the geometric network. After the geometric network was built, the flow direction can be modified with the Set Flow Direction tool. Since the flow direction was verified visually, the tool was run using the With Digitized Direction parameter to define the flow direction. When the geometric network was created, the data was automatically updated to include fields specific to geometric networks, such as the Weights, Enabled, and Ancillary Role fields. These attributes govern the behavior of tracing. The Ancillary Role determines whether the feature is a source or a sink. Within an edit session, the role for vents was set to the sinks and the AC vents were set to the sources. Setting the appropriate role, along with setting the flow direction, provided the ability to run accurate tracing.

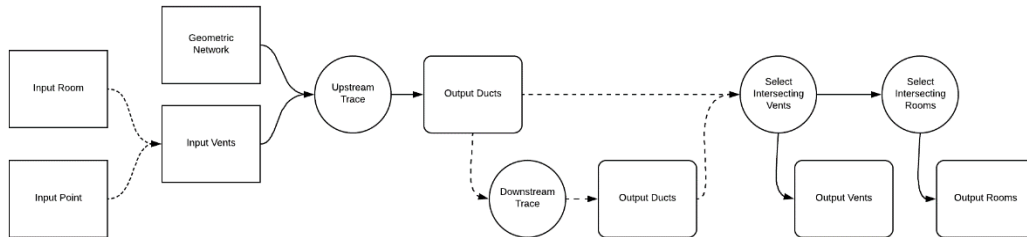
### **5.2.1 Validating the Geometric Network**

Using the Utility Network Analyst toolbar, tracing was performed to validate whether the geometric network was set up correctly. The geometric network was added to ArcMap and the Add Edge Flag Tool was used to place a flag within one of the rooms from the floor plan (Figure 5-12). The Trace Task was set to Find Upstream which would determine the source of a line by tracing against the digitized direction.



### 5.3.1 Building the Geoprocessing Script

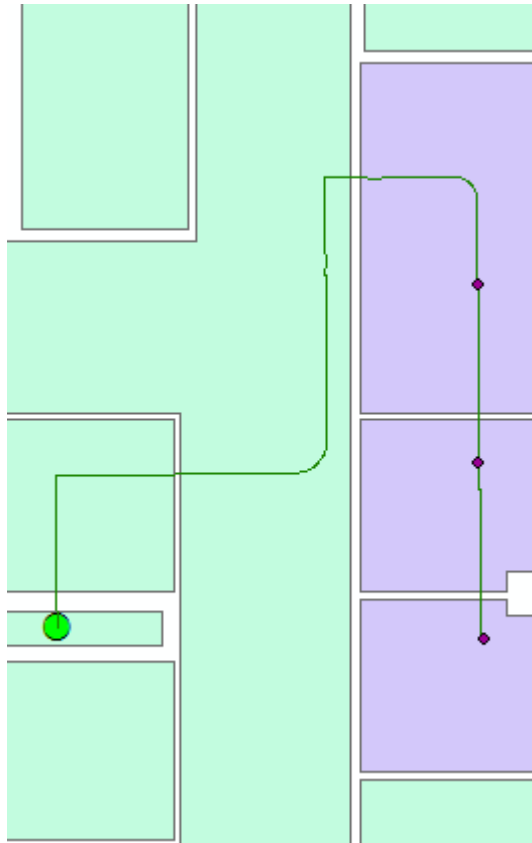
A script was developed to run the analysis through a geoprocessing task. An overview of the design of the tool is presented in Figure 5-13.



**Figure 5-13 Script Tool Design**

There were two separate inputs to the script: a drop down with a list of rooms or a feature set to allow a user to digitize a point. If the user enters a room number, then the Select By Attribute tool is run to select the specific room feature. If a user digitizes a point, then the Select By Location tool is run to select the room feature. In either circumstance, once the room is selected, the Select By Location tool is run to select for the vent that is within the input room. That vent is then used as the input flag for the upstream trace using the geometric network. The result of the trace is the duct features that are upstream towards the source AC unit. The resulting duct features are used to select for intersecting vents and then the rooms that the vents reside in.

The results of the geoprocessing task are three outputs: the ducts returned from the original trace, the vents that intersect the ducts, and the rooms that intersect the vents (Figure 5-14).



**Figure 5-64 Upstream Trace Analysis**

These results can help identify which other rooms are connected to a room where someone contracted a sickness distributed through a utility system. An optional parameter determines whether to run a downstream trace or not. Downstream traces return all features that are with the flow direction, away from the source. If the downstream parameter is selected, the AC unit that intersects the upstream trace analysis is found using the Select By Location tool. The AC unit is used as the input flag for the downstream trace to find all feature that the AC unit is providing air to (Figure 5-15).

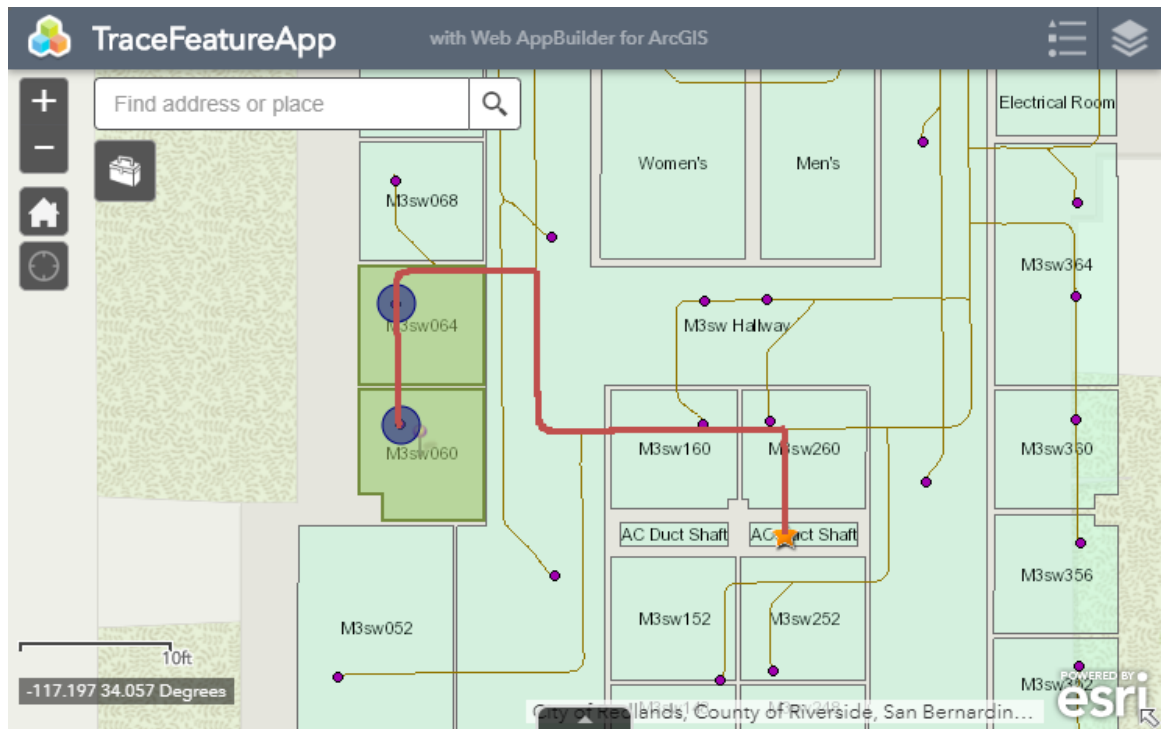


**Figure 5-75 Downstream Trace Analysis Results**

### **5.3.2 Developing the Web Products**

There were three main components involved in developing the web products: the geoprocessing task published as a geoprocessing service, the floor plan published as a map service, and a web application built from the Web Application Builder. Both the services were published to ArcGIS Server and the web application was hosted in ArcGIS Online. If the solution is to be hosted entirely within a customer's environment, Portal for ArcGIS can be used to host the application instead. A geoprocessing widget was added to the web application which was configured to run the geoprocessing service. When the

geoprocessing services completes, it returns the results to the web application as feature layers, seen in Figure 5-16.



**Figure 5-16 Web Application Results**

## 5.4 Summary

There were three main phases to implement the project. The first was converting CAD data to file geodatabase feature classes. Individual tasks included georeferencing the data as well as creating centerlines from the data to act as the ducts. The next major phase was creating the geometric network. Since the flow direction was based on the digitized direction, duct features that were digitized in the incorrect direction needed to be flipped. Lastly, final deliverables were created. They included the geoprocessing script tool to be published as a geoprocessing service and a web application.



## Chapter 6 – Analysis and Results

Chapter 6 discusses the results of the analysis and expands on how the concepts presented in this report can be applied to different settings and problems. It will also cover the challenges with working with CAD data.

### 6.1 Tracing Indoor Utility Features

The tool developed for this project was designed to identify rooms connected to other rooms via a utility system. This can aid in isolating rooms that have been exposed to certain contaminants circulated through an HVAC system or a plumbing system. It can also return the source component of the system that the rooms are connected to, such as an AC unit or hot water tank. HVAC and plumbing systems adhere to the fundamentals of geometric network tracing which include determinate flow and network junctions, so the functionality within the tool will work against both types of utility systems. Given that there were issues acquiring the plumbing system data from the VA Hospital, HVAC data from Esri had to be used. Since the tool is running a trace against a geometric network, the functionality is identical to using water system data from the VA.

The optional parameter of a downstream trace is an important option if the source is assumed to be contaminated. This is particularly useful for analysis of Legionnaires' disease. *Legionella pneumophila* typically colonizes within a water tank and once the bacteria make it into the plumbing lines, it has a chance of turning into an aerosol and being inhaled. By first tracing upstream to determine the source, and then downstream from the source, all rooms connected to the contaminated water tank can be found as depicted by the red rooms in Figure 6-1.

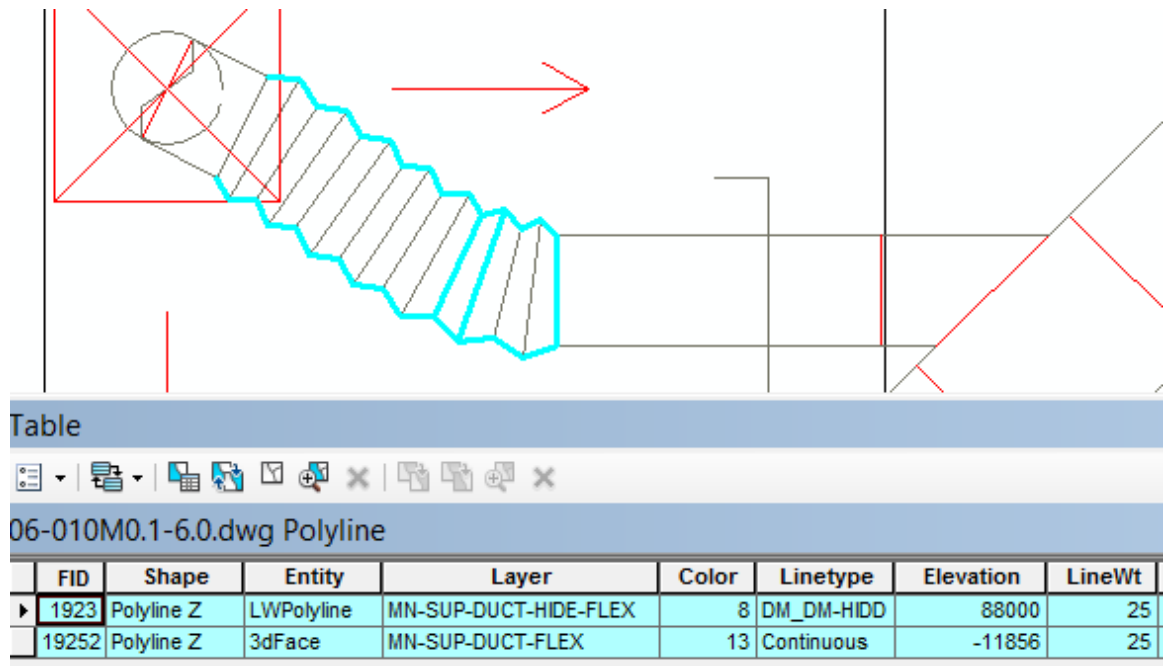




**Figure 6-1 Rooms Identified From Downstream Trace Results**

## **6.2 Converting CAD Data**

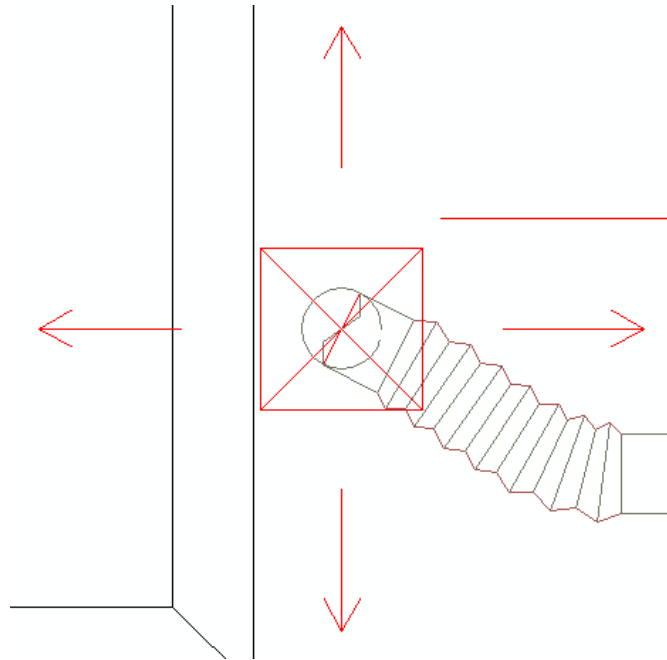
A phase of the project that required significant time was the conversion of CAD data to GIS formats, in this case, feature classes in a file geodatabase. CAD data is not typically drawn with the intention of bringing it into a GIS, so the data model can make it difficult to extract necessary data to convert. Each line type is drawn as a separate layer, and there are multiple different ways a single feature is drawn depending on the line type.



**Figure 6-2 CAD Data Model**

In Figure 6-2, a duct feature was drawn as an individual feature that made up a larger feature, but was also drawn as the larger feature as well (this is represented in the figure as blue lines). Careful SQL statements had to be constructed to ensure that only necessary features were extracted which can be seen in Figure 5-5.

Furthermore, a geometric network requires sinks to be points. However, the CAD data is digitized in a way that vents are line features instead (Figure 6-3).



**Figure 6-3 Selecting for Vents**

It was important to fully discover the data to ensure that necessary values can be selected and extracted to GIS formats. Although generic CAD conversion tools exist, the significant amount of variability in how CAD data is digitized and attributed means that it would be difficult to automate data conversion to the point it can be directly used within something like a geometric network.

## **Chapter 7 – Conclusions and Future Work**

This chapter provides a conclusion of the project and discusses possibilities for future work. Section 7.1 covers the work that was done to complete the project and discusses how the deliverables will be of value to the client. Section 7.2 provides insight into future work in this area.

### **7.1 Conclusion**

The goal of this project was to isolate utility features that were connected to a room with a possible contaminant, such as HVAC ducts and vents, or plumbing features and water outlets. It creates a general framework that can be followed to either identify potential sites of contamination or develop a testing strategy for a facility to use as a preventative method.

The client was the VA Hospital in Loma Linda who have had cases of Legionnaires' disease. Since Legionnaires' disease is spread through the aerosolization of contaminated water particles, the hospital wanted to identify which water lines were contaminated and trace those back to the source water tank. As water line data from the hospital was not available, HVAC data from Esri was used in its place. Both types of systems model network infrastructure so both could be used for this type of analysis.

The first steps were converting the CAD data to file geodatabase feature classes. This involved understanding the data model of the CAD data to extract necessary features. Once the CAD data was converted to feature classes, a geometric network was created. Finally, a Python script was written that ran an upstream trace and used several Select statements, both attribute and spatial queries, to determine vents and AC units that intersected the trace output. Optionally, a downstream trace could be run to return all the

rooms and vents that received air from the AC unit. The outputs were returned as feature layers to the client running the process.

## **7.2 Future Work**

Due to data constraints, data had to be acquired from a different source. Unfortunately, the scale of the HVAC data was far smaller than the hospital floor plans. There are at least two water tanks on each floor of the hospital so a larger data set would result in more complex yet interesting results. An upstream trace would still return a single duct or water line feature, but a downstream trace from the source would return far more features. Since the level of contaminants closer to the source is likely higher than contaminants further away from the source, it may be possible to add weights to the geometric network. This could simulate how much contaminant reached a vent or outlet given the size of pipes leading up to the vent or outlet, and may help in identifying which outlets pose greater risks given their proximity to the source.

While this project developed tools to identify rooms that may already pose risks to patients, the underlying methods can also be used to define testing and mitigation strategies. Consistent testing is an important part of containment efforts. If rooms are tested often, then contamination of the water system can be caught early. An optional parameter within the analysis is a downstream trace from the source. A sampling tool can be constructed that uses a downstream trace to construct a random list of rooms that are connected to a specific source. These rooms would be the ones to test in the next round of sampling. The tool can be part of a consistent monitoring process to contain any contamination within the water system.

One of the more challenging parts of the project was converting CAD data to file geodatabase. A simple CAD to Geodatabase tool does exist, but it converts all CAD features to file geodatabase or enterprise geodatabase feature classes without extracting specific features or feature types. Creating a tool or standardized process that extracts specific CAD features would make conversion to GIS formats easier. If that is done, creating the geometric network could also be automated.



## Works Cited

- Arctur, D., & Zeiler, M. (2004). *Designing Geodatabases - Case Studies in GIS Data Modeling*. Redlands: ESRI.
- Bartram, J., Chartier, Y., Lee, J. V., Pond, K., & Surman-Lee, S. (2007). *Legionella and the prevention of legionellosis*. World Health Organization.
- Beauté, J., Robesyn, E., de Jong, B., Coulombier, D., Hallström, L. P., Takkinen, J., . . . Zucs, P. (2012). *Legionnaires' disease in Europe*. European Centre for Disease Prevention and Control.
- Beer, K. (2015). *Surveillance for Waterborne Disease Outbreaks Associated with Drinking Water — United States, 2011–2012*. Centers for Disease Control and Prevention. Retrieved from <https://www.cdc.gov/mmwr/preview/mmwrhtml/mm6431a2.htm>
- Design Patterns - Linear Networks. (2004). In ESRI, *Building Geodatabases* (pp. 24-25). Redlands, CA: ESRI.
- ESRI. (2018a). *About Spatial Adjustment*. Retrieved from Performing Spatial Adjustments: <http://desktop.arcgis.com/en/arcmap/latest/manage-data/editing-existing-features/about-spatial-adjustment.htm>
- ESRI. (2018b). *What are geometric networks?* Retrieved from Geometric Networks: <http://desktop.arcgis.com/en/arcmap/latest/manage-data/geometric-networks/what-are-geometric-networks-.htm>
- ESRI. (2018c). *About CAD Coordinate Systems*. Retrieved from Elements of CAD Drawing: <http://desktop.arcgis.com/en/arcmap/latest/manage-data/cad/about-cad-coordinate-systems.htm#GUID-1C363358-CBAF-4770-829F-B8752C553C71>



- Fitzgeorge, R. B., Baskerville, A., Broster, M., Hambleton, P., & Dennis, P. J. (1983, February). Aerosol infection of animals with strains of *Legionella pneumophila* of different ferent virulence: comparison with intraperitoneal and intransal routes of infection. *The Journal of Hygiene*, *90*(1), 81-89.
- Garrison, L. E., Kunz, J. M., Cooley, L. A., Moore, M. R., Lucas, C., Schrag, S., . . . Whitney, C. G. (2016). *Vital Signs: Deficiencies in Environmental Control Identified in Outbreaks of Legionnaires' Disease — North America, 2000–2014*. Centers for Disease Control and Prevention. Retrieved from <https://www.cdc.gov/mmwr/volumes/65/wr/mm6522e1.htm>
- Heuner, K., & Swanson, M. (2008). *Legionella - Molecular Biology*. Norfolk, UK: Caister Academic Press.
- Lin, Y.-s. E., Stout, J. E., & Yu, V. L. (2011, February). Controlling *Legionella* in Hospital Drinking Water: An Evidence-Based Review of Disinfection Methods. *Infection Control and Hospital Epidemiology*, *32*(2), 166-173.
- Lin, Y.-s. E., Stout, J. E., Yu, V. L., & Vidic, R. D. (1998, June). Disinfection of Water Distribution Systems for *Legionella*. *Seminars in Respiratory Infections*, *13*(2), 147-159.
- Lin, Y.-s. E., Vidic, R. D., Stout, J. E., & Yu, V. L. (1998, September). *Legionella* in water distribution systems. *American Water Works Association*, *90*(9), 112-121.
- Liu, K. F.-R., Hung, M.-J., Kuo, J.-Y., & Liang, H.-H. (2015). Using GIS and Kriging to Analyze the Spatial Distributions of the Health Risk of Indoor Air Pollution. *Journal of Geoscience and Environment Protection*, 20-25.

- Liu, Z., Stout, J. E., Tedesco, L., Boldin, M., Hwang, C., Diven, W. F., & Yu, V. L. (1994, April). Controlled Evaluation of Copper-Silver Ionization in Eradicating *Legionella pneumophila*. *The Journal of Infectious Diseases*, *169*(4), 918-922.
- Maher, M. M. (2018). *Lining Up Data in ArcGIS*. Redlands: ESRI. Retrieved from Esri Support: <https://support.esri.com/en/technical-article/000012038>
- Mandel, A. S., Sprauer, M. A., Sniadack, D. H., & Ostroff, S. M. (1993, November). State Regulation of Hospital Water Temperature. *Infection Control and Hospital Epidemiology*, *14*(11), 642-645.
- Marrie, T. J., Johnson, W., Tyler, S., Bezanson, G., Haldane, D., Burbridge, S., & Joly, J. (1995). Potable water and nosocomial Legionnaires' disease - check water from all rooms in which patient has stayed. *Epidemiology and Infection*, *114*, 267-276.
- Marston, B., Lipman, H., & Breiman, R. (1994, Nov). Surveillance for Legionnaires' disease. Risk factors for morbidity and mortality. *Arch Intern Med*, 2417-2422.
- McDade, J., Brenner, D., & Bozeman, F. (1979). Legionnaires' disease bacterium isolated in 1947. *Ann Intern Med*, 659-661.
- Orsi, G. B., Vitali, M., Marinelli, M., Ciorba, V., Tufi, D., Del Cimmuto, A., . . . De Giusti, M. (2014). Legionella control in the water system of antiquated hospital buildings by shock and continuous hyperchlorination: 5 years experience. *BMC Infectious Diseases*, *14*(394).
- Paneth, N., & Paul, F. (2013). The singular science of John Snow. *The Lancet*, *381*(9874), 1267-1268.
- Rogers, J., Dowsett, A. B., Dennis, P. J., Lee, J. V., & Keevil, C. W. (1994). Influence of Temperature and Plumbing Material Selection on Biofilm Formation and Growth

- of *Legionella pneumophila* in a Model Potable Water System Containing Complex Microbial Flora. *Applied and Environmental Microbiology*, 1585-1592.
- Services, E. T. (2018). *Working with Geometric Networks for Utilities*. Redlands: ESRI.
- Shiode, N., Shiode, S., Rod-Thatcher, E., Rana, S., & Vinten-Johansen, P. (2015, June). The mortality rates and the space-time patterns of John Snow's cholera epidemic map. *International Journal of Health Geographics*, 14-21.
- Stout, J. E., Lin, Y.-s. E., Goetz, A. M., & Muder, R. R. (1998, December). Controlling *Legionella* in Hospital Water Systems: Experience With The Superheat-And-Flush Method And Copper-Silver Ionization. *Infection Control and Hospital Epidemiology*, 19(12), 911-914.
- Stout, J. E., Muder, R. R., Mietzner, S., Wagener, M. M., Perri, M. B., Deroos, K., . . . Yu, V. L. (2007). Role of Environmental Surveillance in Determining the Risk of Hospital-Acquired Legionellosis: A National Surveillance Study With Clinical Correlations. *Infection Control and Hospital Epidemiology*, 28(7), 818-824.
- Terranova, W., Cohen, M., & Fraser, D. (1978). 1974 outbreak of Legionnaires' Disease diagnosed in 1977. Clinical and epidemiological features. *Lancet*, 122-124.
- Wadowsky, R. M., Yee, R. B., Mezmar, L., Wing, E. J., & Dowling, J. N. (1982, May). Hot Water Systems as Sources of *Legionella pneumophila* in Hospital and Nonhospital Plumbing Fixtures. *Applied and Environmental Microbiology*, 1104-1110.



## Appendix A. Geoprocessing Script for Tracing

```
1  ###
2  ## Name: Trace.py
3  ## Description: This geoprocessing script traces upstream or downstream from input rooms
4  ## either in the form of a value list or digitized point
5  ###
6
7  # Import necessary Python modules
8  import arcpy, os
9
10 = def p(msg):
11     """This function simplifies print statements"""
12     arcpy.AddMessage(msg)
13     print(msg)
14
15 # Overwrite any existing features
16 arcpy.env.overwriteOutput = True
17
18 # Define workspace path
19 workspace = os.path.join(os.path.dirname(__file__), "Esri")
20
21 # Define input parameters
22 inVals = arcpy.GetParameterAsText(0)
23 inFeat = arcpy.GetParameterAsText(1)
24 runDownstream = arcpy.GetParameter(2)
25
26 # Construct paths to data
27 ventsPath = os.path.join(workspace, r"Data.gdb\GN\Vents")
28 floorPlanPath = os.path.join(workspace, r"Data.gdb\MFloorPlan")
29 inGN = os.path.join(workspace, r"Data.gdb\GN\HVAC")
30 ductsPath = os.path.join(workspace, r"Data.gdb\GN\Ducts")
31 ventsFL = arcpy.MakeFeatureLayer_management(ventsPath)
```

Figure A 1 Trace Features (Lines 1-31)

```

31 ventsFL = arcpy.MakeFeatureLayer_management(ventsPath)
32
33 # Check whether a value from the value list was selected
34 # or an input point was digitized. If neither, return an error
35 = if inVals:
36     ventsLayer = arcpy.MakeFeatureLayer_management(ventsPath, where_clause="RoomNumber = '{}'".format(inVals))
37 = elif int(arcpy.GetCount_management(inFeat)[0]) > 0:
38     roomsFL = arcpy.MakeFeatureLayer_management(floorPlanPath)
39     arcpy.SelectLayerByLocation_management(roomsFL,"INTERSECT",inFeat)
40     ventsLayer = arcpy.SelectLayerByLocation_management(ventsFL,"WITHIN",roomsFL)
41 = else:
42     arcpy.ExecuteError("Enter a room number or place a point within a room.".format(inVals))
43
44 = try:
45     # Check whether there are features in the vents layer.
46     # If not, return an error.
47 = if int(arcpy.GetCount_management(ventsLayer)[0]) == 0:
48 =     if int(arcpy.GetCount_management(inFeat)[0]) > 0:
49         rooms = []
50         with arcpy.da.SearchCursor(roomsFL,['LONGNAME']) as cursor:
51             for row in cursor:
52                 rooms.append(row[0])
53             inVals = ", ".join(rooms)
54 =             if len(rooms) > 1:
55                 msg = "No vents found within rooms {}".format(inVals)
56 =             else:
57                 msg = "No vents found within room {}".format(inVals)
58             arcpy.AddError(msg)
59 =     else:
60         # Run upstream geometric trace using the vents as the input flag
61         outGN = arcpy.TraceGeometricNetwork_management(inGN,"",ventsLayer,"TRACE_UPSTREAM")
62         p("Ran upstream trace.")

```

**Figure A 2 Trace Features (Lines 31-62)**

```

62     p("Ran upstream trace.")
63
64     # Create a layer of the duct features returned by the trace
65     resultDucts = arcpy.MakeFeatureLayer_management("{}\\Ducts".format(outGN))
66
67     # Run a downstream trace after finding the AC Unit that intersects the trace result
68 = if runDownstream == True:
69     acVents = os.path.join(workspace,r"Data.gdb\GN\AC_Unit")
70     acFL = arcpy.MakeFeatureLayer_management(acVents)
71     arcpy.SelectLayerByLocation_management(acFL,"INTERSECT",resultDucts)
72     outGN = arcpy.TraceGeometricNetwork_management(inGN,"",acFL,"TRACE_DOWNSTREAM")
73     p("Ran downstream trace.")
74     resultDucts = arcpy.MakeFeatureLayer_management("{}\\Ducts".format(outGN))
75
76     # Create a layer of the vents features returned by the trace
77     resultsVents = arcpy.MakeFeatureLayer_management("{}\\Vents".format(outGN))
78
79     # Select the rooms that intersect the ducts
80     roomFL = arcpy.MakeFeatureLayer_management(floorPlanPath)
81     arcpy.SelectLayerByLocation_management(roomFL,"INTERSECT",resultsVents)
82
83     # Remove the hallway and duct shafts from the rooms
84     arcpy.SelectLayerByAttribute_management(roomFL, "REMOVE_FROM_SELECTION", "LONGNAME in ('M3sw Hallway','AC Duct Shaft')")
85     resultsRooms = arcpy.MakeFeatureLayer_management(roomFL,"Rooms")
86     p("Determined intersecting rooms.")
87
88     # Return the ducts, vents, and rooms as output features
89     arcpy.SetParameter(3,resultDucts)
90     arcpy.SetParameter(4,resultsVents)
91     arcpy.SetParameter(5,resultsRooms)
92 = except Exception as e:
93     arcpy.AddError(e)

```

**Figure A 3 Trace Features (Lines 62-93)**