Task 2, Section A

FS InfoCat – File System Information Cataloging Tool

Leonard T. Erwine

Western Governors University

000356334

Table of Contents

[The Business Problem 3](#_Toc83250290)

[The Customer 3](#_Toc83250291)

[The Business Case 4](#_Toc83250292)

[Fulfillment 5](#_Toc83250293)

[Existing Gaps 5](#_Toc83250294)

[Software Development Lifecycle 6](#_Toc83250295)

[Requirements Validation 6](#_Toc83250296)

[Planning and Design 6](#_Toc83250297)

[Deliverables 7](#_Toc83250298)

[Deliverables for Pastoral Staff and Stakeholders 7](#_Toc83250299)

[Technical Deliverables 7](#_Toc83250300)

[End-User Deliverables 7](#_Toc83250301)

[Implementation Phase 8](#_Toc83250302)

[Software Implementation and Concurrent Testing 8](#_Toc83250303)

[User Acceptance Testing 8](#_Toc83250304)

[Programming Environment 8](#_Toc83250305)

[Development Environment and Cost 8](#_Toc83250306)

[Other Costs 9](#_Toc83250307)

[Project Timeline 9](#_Toc83250308)

# The Business Problem

## The Customer

Centreville Baptist Church is active with an extensive community support ministry whose staffing predominantly comprises part-time and volunteer positions. In addition to the humanitarian services provided to the community, the audio/visual team embraces the newest technologies, making full use of the church’s IT assets. The services provided by volunteer personnel tend to be transient but still very valuable to the success of the overall goals of the local religious community. As technology plays an increasingly vital to even the most menial of tasks, the support staff is continually working to ensure that IT resources are properly maintained and available to all participants. Inevitably due to the part-time nature of the work schedules, there is not always enough time to keep up with everything. In addition, as systems degrade due to inattention, IT staff frequently finds themselves reacting to incidents rather than proactively preventing them. The management of storage devices and their content is one area that often presents problems. Many times, the A/V team would like to re-use material that someone had previously created. However, they cannot always locate the files because the original creators are not available for consultation. Additionally, the employment of new technology increases the demand for storage space, leading to the depletion of free space on shared drives.

The pastoral staff worked with the IT staff to formulate policy and governance on the usage of storage devices but found it cumbersome to verify or enforce due to the transiency of most participants. They determined that instituting an IT training regimen would not be beneficial in most instances. If volunteers were required to complete training sessions before utilizing IT equipment, it would detract from the time spent on their humanitarian work, which is the church’s primary goal.

## The Business Case

FS InfoCat will scan files distributed across multiple storage locations, devices, and media, creating a catalog of characteristics about each discovered file. The information in the database can then be searched and analyzed to look for specific patterns or values which may expose unwanted files. More specifically, this software will assist with the mitigation of shared network drive overuse by identifying possible file duplication or files with a size that has exceeded a specific threshold. Not only does it compare by file size, but it also calculates cryptographic hash codes of files that share the same length. Thus, files with the same size and cryptographic hash are highly likely to be exact copies of one another. Additionally, the software offers a way to view file metadata according to their file type, allowing users to identify redundancies even if the file sizes differ. Finally, the user can exempt files and subdirectories when redundancy is permitted or notate them for further action.

A pastoral staff member initially objected to the prospect of using the software on local workstation hard drives, citing the risk that confidential information could be compromised if metadata gets published to the shared database. This risk was mitigated by separating the data gathering process from the publication process. The application will initially store information to a local database within the user’s windows profile. It will only get published to the centralized database when the user explicitly opts to do so. More specifically, while the software application can share information with a centralized database, it will not depend upon it and function independently. Additionally, users can configure the software to skip specific files and folders, and they can specify multiple configurations that only scan particular folders.

## Fulfillment

FS InfoCat will be a windows desktop application. It will leverage the Windows Property System to extract metadata from files, such as author names, audio titles, and video frame rates. Since not all devices support the Windows Property System, the application allows people to provide relevant information for each file manually.

# Existing Gaps

Shared network drives used by the staff and volunteers alike lack sufficient structure. The cost of file storage steadily increased, but the value to the business was not proportionate due to the wasted space. There were many instances where a staff member had to re-create something simply because nobody knew where the person who did the original work saved all their files. They also found themselves periodically running out of storage space, causing work stoppage or lost data. Manually cleaning up the common storage areas may be as costly as the price of increasing storage capacity. Due to the uncertainty of the actual content within the shared drives, it is impossible to assess the cost in person-hours, should they decide to go through each file to determine what can be consolidated or discarded. It is also conceivable that the pace of unmanaged data contributions could outpace the mitigation efforts. Even if they did reduce the storage footprint, reducing the cost of file storage services will not solve the problem of recalling specific stored files when needed.

The original intent was to produce an application in the first iteration that could synchronize with a centralized database. However, in the interest of utilizing the software sooner, the first release will not include this capability and will be deferred to the next release. However, in this state, it is still viable for gathering and analyzing information locally.

# Software Development Lifecycle

I will design this software using a modified Waterfall development approach. The main reason for using this method is that there will most likely only be a single developer. Another factor contributing to this decision was that none of the stakeholders think they will consistently have enough free time to participate in Scrum meetings. Additionally, the capabilities required in the first release would require much more time than I could fit within a typical Agile sprint.

## Requirements Validation

The development process will begin with validating the requirements gathered in the initial meeting with the customer. In doing so, I will explicitly state all assumptions to be sure that I understand customer expectations. Then, to capitalize on the advantages of the Agile disciplines, I will organize the use cases and requirements in a GitHub project. From there, I will define the structure for the data models. Then, using several detailed diagrams, I will depict the design of the user interface to include how each part of the user interface corresponds with database entities.

## Planning and Design

Next, I will define the software validation process using three stages: Automated testing, scripted developer testing, and user acceptance testing. Automated unit tests are created synchronously with the development process to validate each component as I implement them and expose potential bugs. The purpose of the scripted developer testing is to ensure that it is fit for use. I will script the tests to ensure all features within the graphical user interface produce the intended result. Finally, the church’s system administrator and some audio/visual team stakeholders will participate in the user acceptance testing. The user acceptance testing will ensure that the software meets its intended purpose as understood by the customer. I will submit the entity model, user interface design, and testing plans to the customer for final scrutiny before I begin the implementation phase.

# Deliverables

## Deliverables for Pastoral Staff and Stakeholders

Before initiating the implementation phase, I will prepare a requirements document that enumerates the use cases, expectations, and parameters set forth during the requirements gathering phase. Additionally, I include the entity models, user interface diagrams, and testing plans created during the design process. Finally, each test will need to indicate a way to measure performance or a metric of success.

## Technical Deliverables

Upon completing the implementation phase, I will finalize technical instructions on deploying, configuring, and monitoring the software. The system administrator will validate these technical instructions during the first part of the user acceptance testing phase. Then, at the start of the user acceptance testing, I will deliver the technical documentation and the software binaries to the system administrator.

## End-User Deliverables

I will create documentation that will be accessible to the end-user on the features and intended usage of the software. The user documentation can be a mixture of in-application documentation and tooltips, as well as online documentation. I will also produce a documentation file that contains any information presented in online documentation as a fallback when online resources are unavailable.

# Implementation Phase

I will use Visual Studio 2019 as the development platform because it is better suited to using Windows operating system APIs than other platforms. This way, by using an OS-native platform, I will decrease the likelihood of unforeseen complications.

## Software Implementation and Concurrent Testing

During the implementation phase, I will use test-driven development techniques for the components that are critical to reliability or security. Additionally, code segments that are difficult to monitor or validate using event tracing or other debugging means will be validated using unit tests. For the test-driven components, I will write the tests before implementing the code. After the code has passed all unit tests, I will configure the unit tests to automatically run with every build, to ensure additional changes produce no adverse effects. Otherwise, I will carry out the scripted developer testing after implementing each target user interface element. I will also re-execute scripted developer tests at the culmination of the implementation phase.

## User Acceptance Testing

The final testing phase will have two parts. The first part will be conducted with the system administrator to ensure that it can be installed and executed in the target environment. Upon the second part of the testing phase, we will install the software on the workstations of designated users who will evaluate the software. Users will have a predetermined series of actions to invoke within the application and verify the expected result.

# Programming Environment

## Development Environment and Cost

Vs

## Other Costs

$$

# Project Timeline

\_\_

References