Lab 1 – Blackboard Archive Extractor Description

Dominic Nguyen

CS 411

Professor Thomas J. Kennedy

April 22, 2017

Table of Contents

List of Figures

List of Tables

1. Introduction

According to Corcoran (2014), approximately 17,000 schools were using Blackboard, in over 100 different countries. Blackboard is an online tool used by schools, colleges, and universities to organize course materials[1]. This tool allows professors to distribute course materials such as syllabi, notes, and assignments to students over the Internet.

According to the Old Dominion University Faculty Handbook, lecturers are reviewed every year, senior lecturers are reviewed every three years, and tenured professors are reviewed every five years[2]. Depending on one's status with the University, lecturers rely on Blackboard to create their portfolio that will then be evaluated. According to the ODU policy, courses on Blackboard are limited to only storing courses with a maximum of 2 years. Professors are forced to either move their content to an empty course or export their course as a Blackboard archive file. A tenured professor would have to move their materials at least twice between evaluations. If the lecturer opted to create a Blackboard archive of their course, they would have to manually extract their content page-by-page. They would also have to manually format the folder structure to allow it to be viewable. This process would have to be repeated for each course archive. Professors have difficulty retrieving their course content after being exported to a Blackboard Course Archive.

Blackboard Archive Extractor is an application designed to allow a user to load a Blackboard course archive and explore its contents. The application will be able to handle selected Blackboard content such as blogs, modules, documents, and exams. The

application will then create a directory containing all the extracted and processed contents from the Blackboard Course Archive.

2. Blackboard Extractor Product Description

Blackboard Archive Extractor application will consist of major components including a user interface, algorithms to recreate the course, and an output that will be reviewable. The application will be desktop software available for Windows 7 through Windows 10. The user will then have the option to choose where to save their reformatted Blackboard file. Upon completion, the application will create a navigable set of HTML documents that will display a webpage with the user's materials from a Blackboard course.

The Blackboard Archive Extractor will go through various algorithms and processes, which include linking and formatting. The linking algorithm will follow the process as in Figure 1. The Blackboard Archive Extractor will take in the Blackboard archive and parse through the archive for manifest files. It will find the associated resource files and link them to the root. The algorithm will then find the next child of current resource file and link it to the parent.
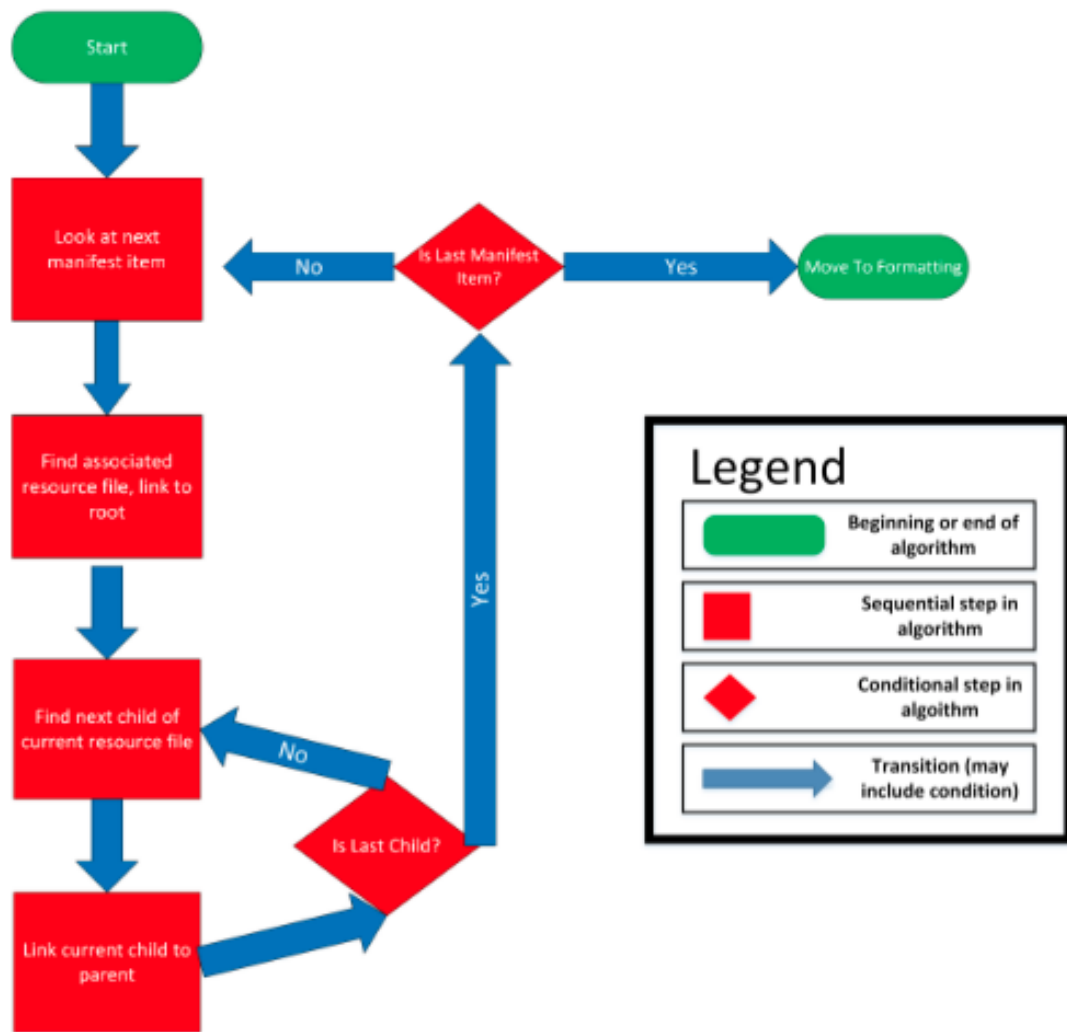
Figure 1: Linking Algorithm

These children indexed may vary from journal post to homework assignment and lead

into their respective resources as depicted in Figure 2. This process will continue until the

last child is linked and continuing until every manifest file has been utilized. Once this is

done, the linking algorithm will end and the formatting algorithm will begin. The

formatting algorithm will follow the process specified in Figure 3.

Figure 2: Linking Algorithm – Resulting Data Structure

Figure 3: Formatting Algorithm

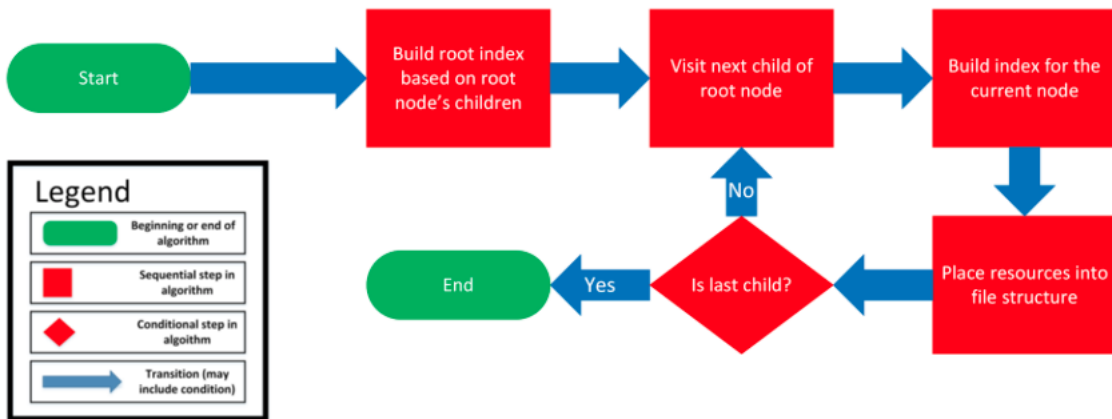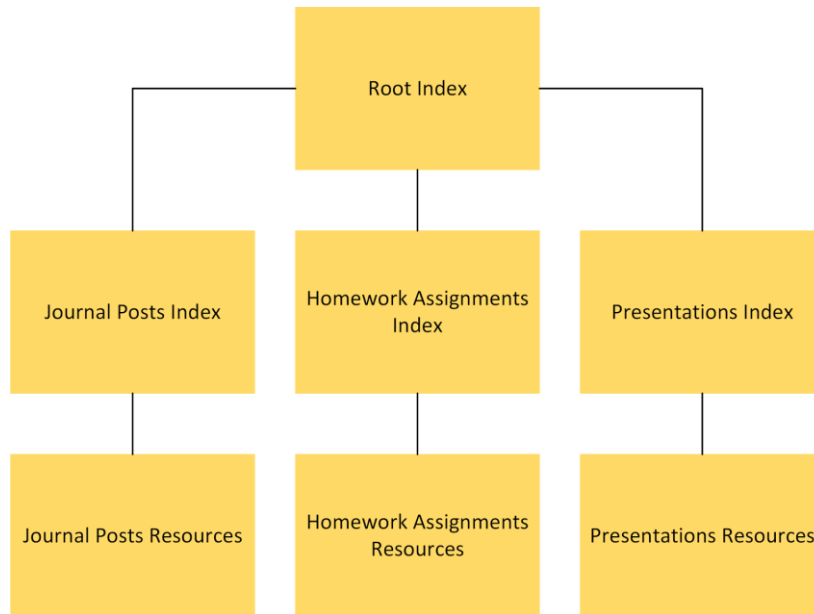The algorithm will build a root index based on the root node's children. It then visits the

next child of the root node and builds an index for the current node. From there, it will

place the resources into a file structure until the last child is accounted for and the process

will then be completed.

   The Blackboard Archive Extractor will then create an output folder for the user to

access. This folder can be saved to the hard disk. It will also create a web content from

the folder and have a responsive website design to host the users course. The materials

will be accessible as they would be on Blackboard, however with the Blackboard Archive

Extractor, the content will be viewable for the Evaluation Committee to review.

   The Blackboard Archive Extractor will consist three major components, these

including parsing, linking, and reformatting. The major functional components are as

specified in Figure 4. The structure Blackboard Archive consists of a manifest file. The

manifest file contains IDs that reference other files that make up various contents of the
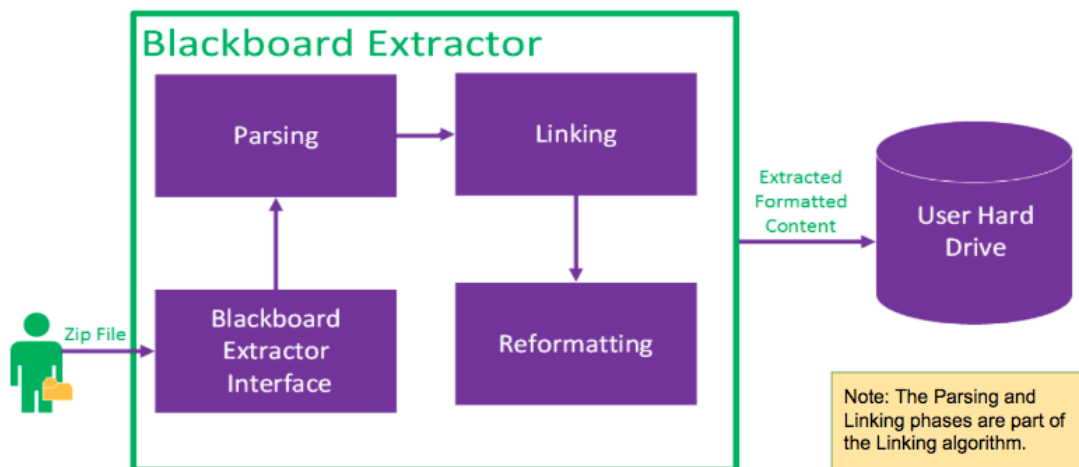
course as shown in Figure 5.

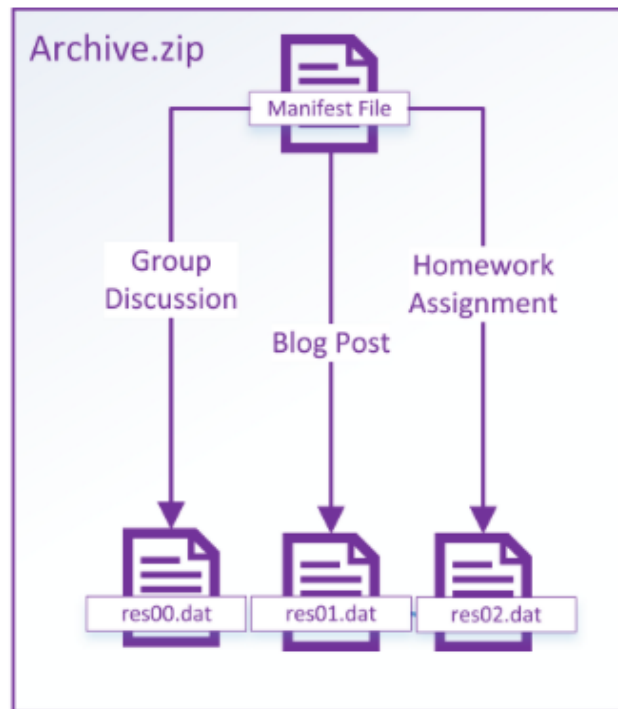**Figure 4: Major Functional Component**

Figure 5: Structure of Blackboard Archive

These subsequent files may be pointing to .dat file, which may entail a group discussion, blog post, homework assignment, or various materials uploaded in that course. This file is difficult to decipher, which is why the Blackboard Archive Extractor is needed.

The archive which contains these manifests and subsequent files is uploaded into the Blackboard Archive Extractor interface and begins the parsing phase. In this phase, the software goes through the archive and creates a directory tree that begins at the root. The parsing will continue until every file is associated with a resource file.

This then leads into the linking phase. In this phase, the Blackboard Archive Extractor will cycle through the entire archive and link every child to its parent. Once this occurs and all the manifests have been read, the software will undergo the reformatting phase.

In the reformatting phase, the application will recreate file structure, reformat HTML, and be viewable. The software will build a file structure containing all the contents that were parsed and will be indexed by HTML files. These files will be accessible links. At the end of the reformatting phase, the entire structure of the recreated archive will be viewable in the set of HTML pages and can be hosted as a webpage with all the materials that were hosted on the Blackboard course.

With the wide use of Blackboard across the world, any user can adopt the Blackboard Archive Extractor. The target customer base is Old Dominion University (ODU) faculty and all educators using Blackboard. ODU faculty will be the primary target in which the software will be implemented. They will be vital to the beta testing. It would be an open testing environment conducted by ODU faculty to focus on the formatting and display for the users. The Blackboard Archive Extractor will require significant acknowledgement to be adopted by other educators from different areas. The Blackboard Archive Extractor will be fully functional to be utilized by all consumers for purposes varying to the user. The Blackboard Archive Extractor will start locally and gradually expand to other schools, colleges, and universities.

[This space intentionally left blank]

3. Blackboard Extractor Prototype

The Blackboard Archive Extractor will implement every major functional component

and algorithms required for the final product as illustrated in Figure 6.

Features	Real World Product	Prototype
Can extract all content from an archive	Fully Functional	Partially Functional
Saves archive on hard disk	Fully Functional	Fully Functional
Checks every URL that points to an outside website	Fully Functional	Partially Functional
Execute application from command line interface	Fully Functional	Fully Functional
Execute program from a GUI	Fully Functional	Fully Functional
Doesn't break on future Blackboard versions	Fully Functional	Partially Functional
Executes from a web application	Fully Functional	Eliminated
Creates web content from archive	Fully Functional	Fully Functional
Responsive website design	Fully Functional	Partially Functional

Table 1: Real World Product with Prototype

[This space intentionally left blank]

The application will be distributed through download from the website. The software will extract all content from the Blackboard archive. The prototype will check every uniform resource locator (URL) that points to an outside website. It will save the new archive onto the hard drive in the specified location. The Blackboard Archive Extractor will execute the program from both command line interface (CLI) and/or graphical user interface. It will also create a web content form the archive and have a responsive website design. Lastly, this prototype will be compatible with future Blackboard versions.

The prototype will require mitigations to various technical and customer challenges. Blackboard abandons IMS Course Cartridges and modifies archive formats are some difficulties that may require mitigation.

Blackboard may undergo updates that will alter the archive format and eliminate Course Cartridges format, the software would have to be updated. It would be good practice to support the latest version of Blackboard and to utilize standard object oriented design principles. Also, Blackboard may update the Blackboard Archive format within the next two years. This would cause the Blackboard Archive Extractor to lose its functionality. To mitigate this, it would be ideal to have good software development practices, which will all the extractor to be updated with ease when necessary.

Customers may not like the user interface (UI) and user experience (UX), they could be hesitant to use the software. To overcome this, there will be multiple interface formats offered and allow users to voice their opinions.

4. Glossary

Blackboard Learn – is a virtual learning environment

Blackboard Archive – a ZIP file of the course content and user interactions

Blackboard Archive Extractor – a tool that extracts blackboard archives into an easily navigable set of HTML documents

GUI – Graphical User Interface is a type of user interface that allows users to interact visually

IMS Course Cartridge – is a specification that describes format for creating and sharing primarily educational digital content

Old Dominion University (ODU) – is a public co-educational research university located in Norfolk, Virginia

Root Index – where the website files are stored

UI – User Interface is where a user and system interact

URL – Uniform Resource Locator a web address referencing a web resource, specifying location

UX – User Experience is the overall experience when using a product

ZIP file – is an archive file format that supports lossless data compression

5. References

1. Corcoran, B. (2014, July 23). Blackboard's Jay Bhatt Strikes Up the Brass Band. Retrieved October 19, 2016, from https://www.edsurge.com/news/2014-07-23-blackboard-s-jay-bhatt-strikes-up-the-brass-band

2. ODU Faculty Handbook. (2005, December 21). Retrieved October 26, 2016, from http://ww2.odu.edu/ao/facultyhandbook/index.php?page=ch07s01.html