**5-2 Milestone Four: Enhancement Three: Databases**

Tomas J. Estanislao

Southern New Hampshire University

CS 499: Computer Science Capstone

Instructor Maryann Krupa

August 24, 2024

**5-2 Milestone Four: Enhancement Three: Databases**

Part 1:

1. Briefly describe the artifact. What is it? When was it created?

In contrast to my previous enhancements, I am focusing on the original web scraper code— which provided the database that I have been working with. I built the scraper using python’s BeautifulSoup framework, while also following along with tutorials and guides. It was created a few months ago and was made to scrap web novel titles. Sourcing from off the publicly available website royalroad.com.

1. Justify the inclusion of the artifact in your ePortfolio. Why did you select this item? What specific components of the artifact showcase your skills and abilities in Databases? How was the artifact improved?

I selected this item because I have been wanting to explore more about data collection and manipulation. I enjoy reading insights on data and manually gathering my own datasets, rather than acquiring a ready-made set. This artifact demonstrates my abilities in gathering and processing data, while also showing how I can interact with databases to provide solutions. I have been able to combine the insights gained from previous enhancements, together with my initial plan, in order to deliver a useful API and dataset. I made some significant improvements which include: re-doing my termination conditions, protecting the storage file, processing the data by converting the data types, and implementing an automatic seeding program. I was able to test the functionality of the changes by comparing the before and after of my enhancement two API querying feature.

1. Did you meet the course objectives you planned to meet with this enhancement in Module One? Do you have any updates to your outcome-coverage plans?

Initially I wanted to refactor the SPA from the original full stack project. However, I decided that focusing more on the backend would allow for me to explore work that may better align with my career goals. Nonetheless, I believe I met the objectives of using appropriate solutions to gather data, while also accomplishing the goal of making that data manipulatable. I even touched on smaller aspects of security, such as appropriately guarding protected flies to avoid unwanted change.

1. Reflect on the process of enhancing and modifying the artifact. What did you learn as you were creating it and improving it? What challenges did you face?

I learned that there are multiple ways to solve even the simplest of tasks. I found that it was better to not overthink how I want something to be done, but rather where I want the project to go. Especially during a time in my life where I am meant to be learning new things and experimenting— it’s important to try things. I was conscious on not using the first piece of code that popped up when searching the internet, while making sure that I understood what each new line does. An example of where this can go wrong is “Spaghetti Code”. I investigated how others solved similar problems with data cleaning and checking files. I tested with numerous print statements and changed the boundaries to test on a smaller scale. I revisited my previous enhancement in order to test my new dataset’s usability.

I completed my enhancement on my original repository. All changes are within the **mainScraper.py** and **seedDB.py** files in the “**Enhance3”** branch. I zipped and published onto my ePortfolio github repo.

Web Scraper Repository: <https://github.com/Testanis396/WebScraper/tree/Enhance3>

ePortfolio Repository: <https://github.com/Testanis396/Testanis396.github.io/tree/main>

Post-Enhancement Review Video: <https://youtu.be/r_Ut86hPLD4>