LTU-RKive.org: Developing a Web Application to Archive Research Artifacts for LTU Research Day and the Library

Evan Wheeler, BSCS Candidate, CoAS; Derek D'Arcy, BSCS Candidate, CoAS; CJ Chung, PHD, Professor, CoAS

College of Arts and Sciences, Lawrence Technological University

Abstract

The LTU-RKive project aims to be an LTU hosted database-enabled web application server for students and faculty to upload, store, browse and search for their research artifacts including annual research day poster presentations, master theses, PHD dissertations and additional materials in PDF file format. This project includes secure user accounts enabling upload ing of their pdf files, after those files are checked for viruses, the ability to interact with a robust search interface, login and account management, and document hosting with content management. The project was also set out with the ability to validate contents of files, however this feature was cut for time. Current testing results show that the developed system satisfies the planned requirements, and as such the website is live at https://ltu-rkive.org, and is ready for archiving Research Day materials, as well as any other artifacts for the LTU Library.

Introduction

Currently, at LTU there is a Research Day hosted every year, which gives students and faculty members the opportunity to present their work from the year. However, once these works are presented at Research Day, their posters have not been archived and shared. Other research artifacts such as Master Theses and PhD dissertations are currently archived in the paper format only in the school library. Through conducting surveys as shown in a section below, targeting students and faculty who are giving presentations at research day, we discovered that the large majority of those who submit these Research Day works wish that there was a better way to collect, store, update, and share their work. We could not use popular open-access archiving site such as Cornell University's arXiv[1] because they do not support poster format. Research Day posters, presentations are the culmination of their hard work, and as such deserve a location to store, present and share the work. This site gives students, staff, and the library the ability to showcase the work and access groundbreaking research from anywhere, and promote ease of access to everyone, regardless of their LTU student status. This site is currently live, and ready to host the masterful works submitted by LTU affiliated researchers!
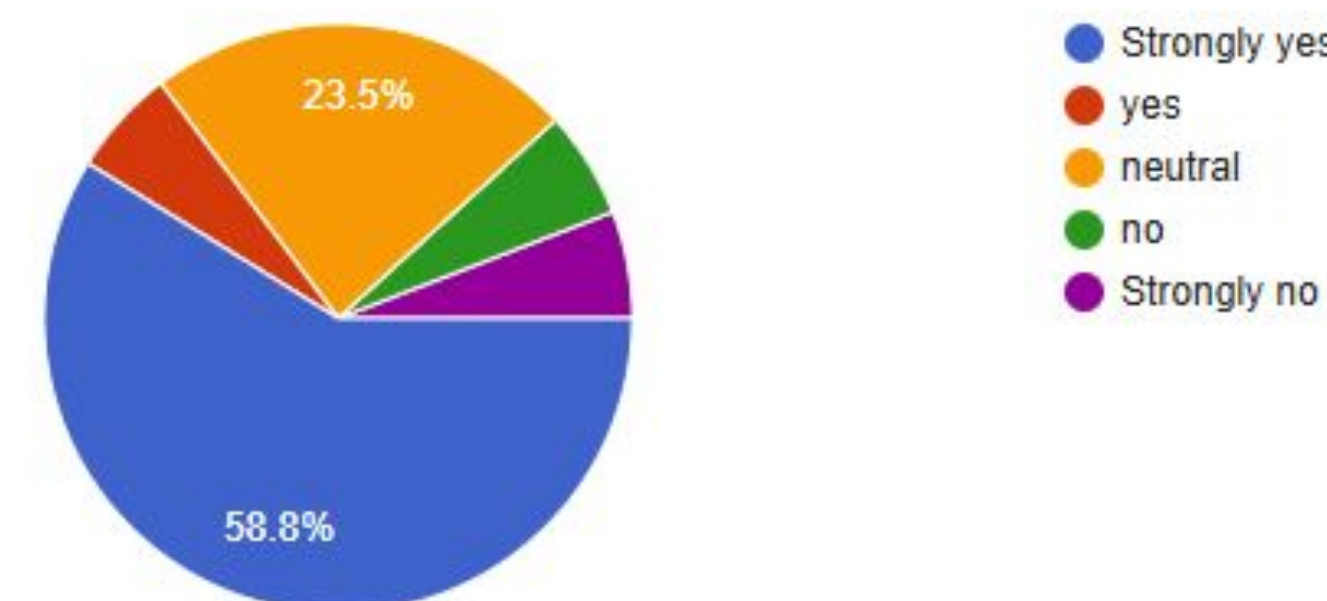
Survey Results

Q2. Would you like your posters in PDF form accessible by yourself and others on the web?
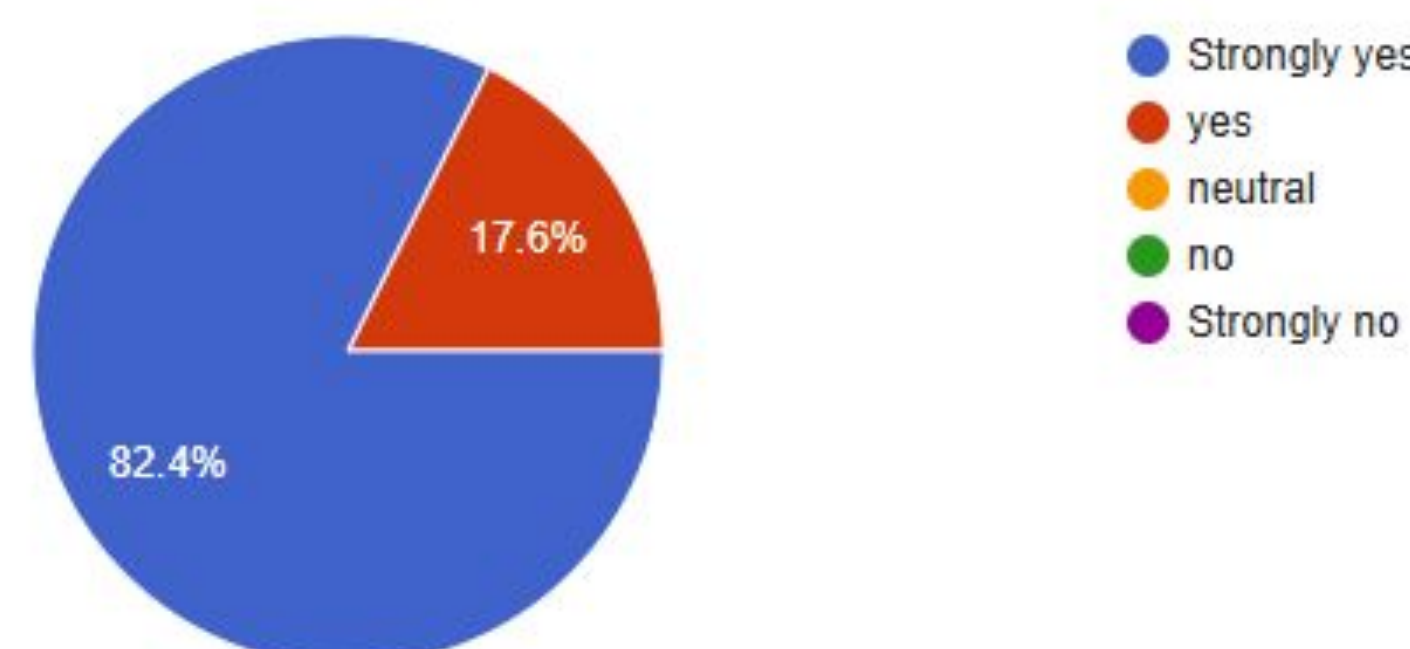17 responses
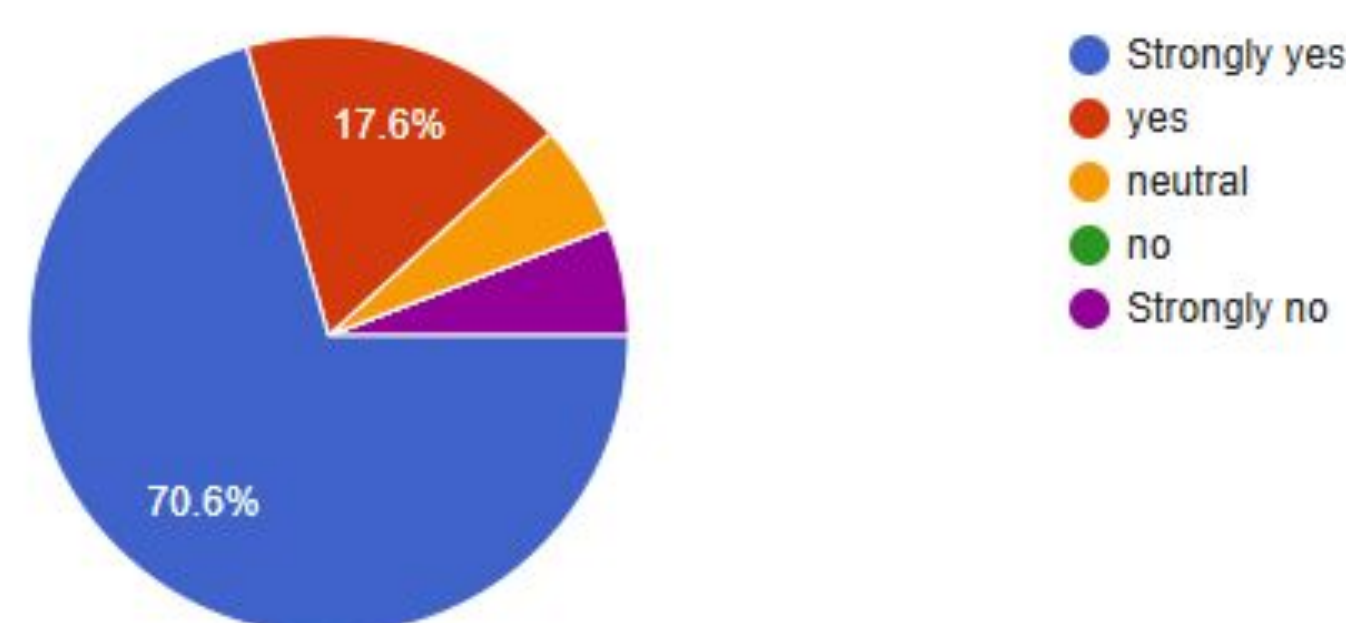
Q3. Would you like to have the link (URL) of your poster posted on your social media such as LinkedIn?
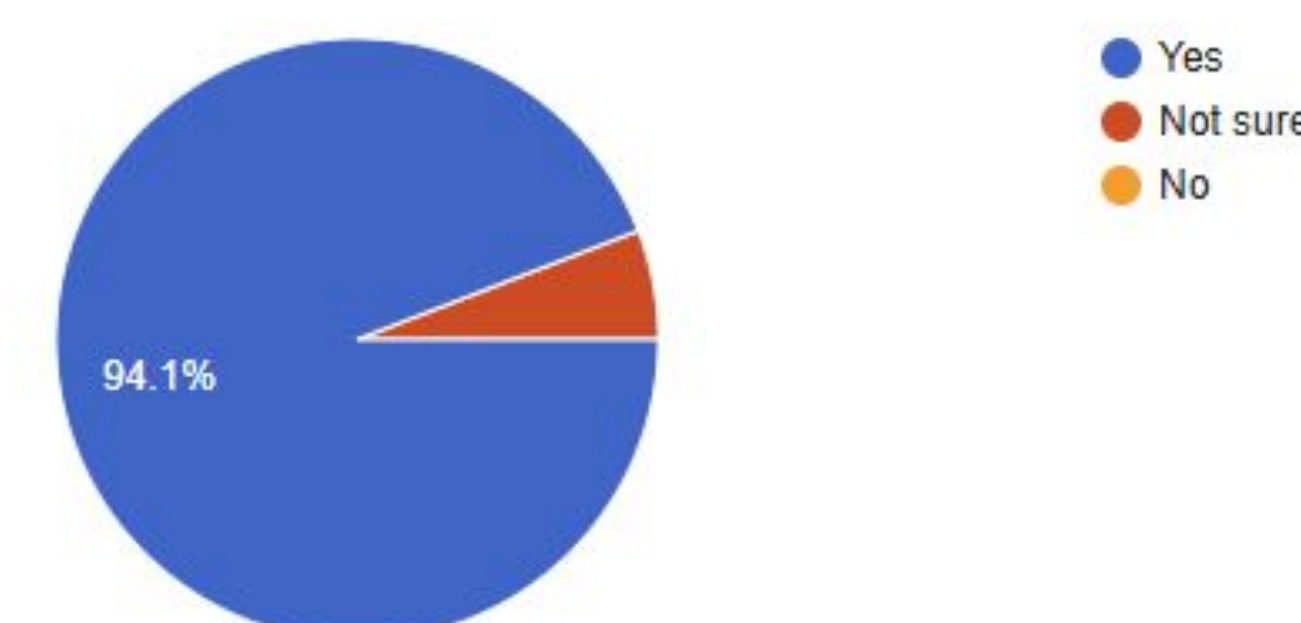17 responses

Q4. Dr. Chung and his Sr Project students are developing an website to archive research artifacts such as posters, presentation slides, technical reports, theses, and dissertations. Would you like to use the system?
17 responses

Q5. Do you believe the archiving system could become part of the LTU library's services?
17 responses

Design

The technical specs to implement this utilize GOlang[2] as the main driving code. Go was chosen as the language for this project due to our familiarity with the language, it's simplicity and speed, it's built in tools and the overall performance. HTML and CSS are used for the front end of the website, with SQLite serving as the database foundation Originally, an LTU hosted server running Ubuntu hosted the project. Due to difficulties with the VPN and security concerns, the server has since been moved to an AWS instance that will now host the site. In addition, there is an email server set up allowing for seamless email integration, which allows users who forgot their passwords to request an email with a password reset. The site also utilizes a Cloudmersive API[3] integration, which serves as the validation and scanning for viruses, ensuring the safety and security of the site and backend database.
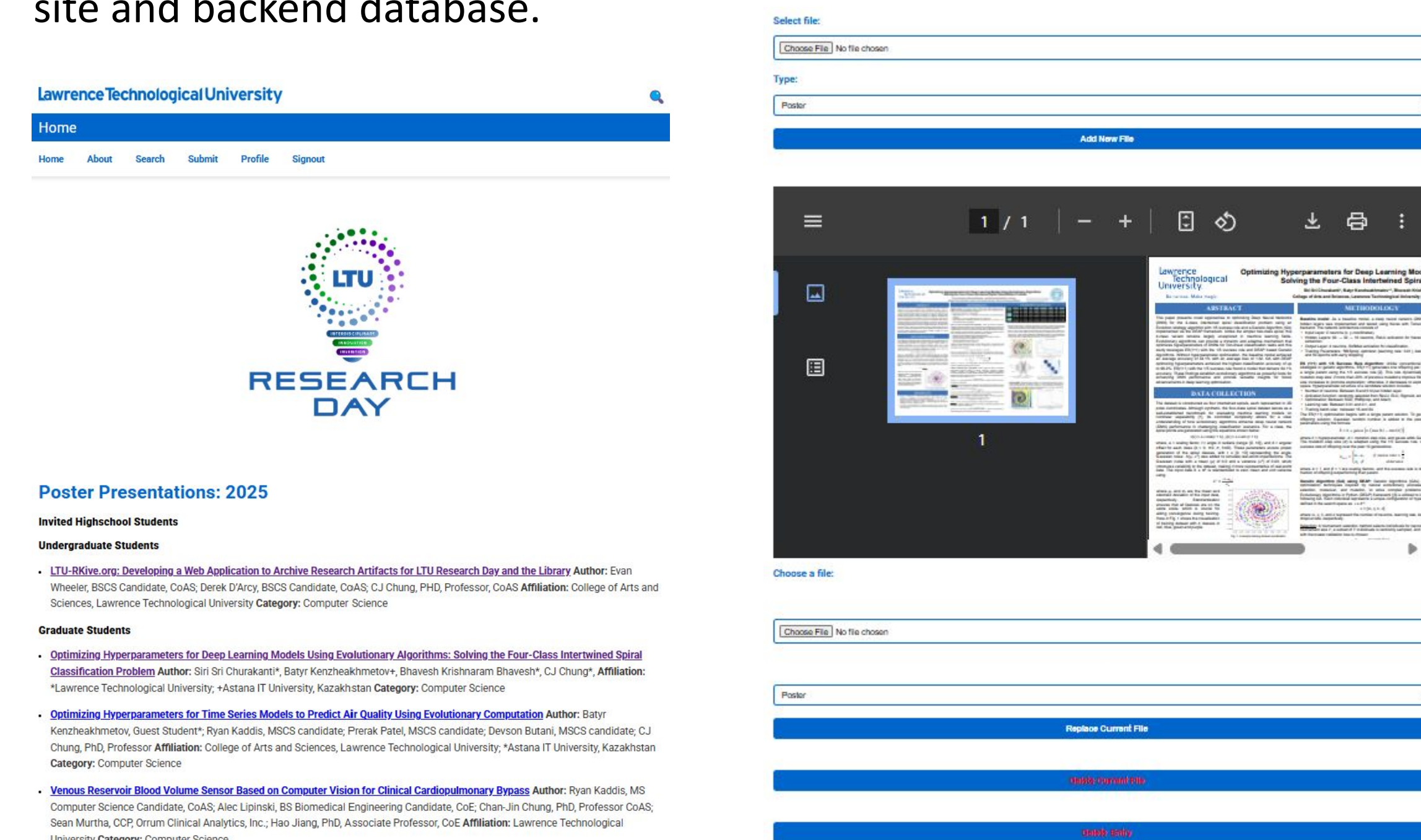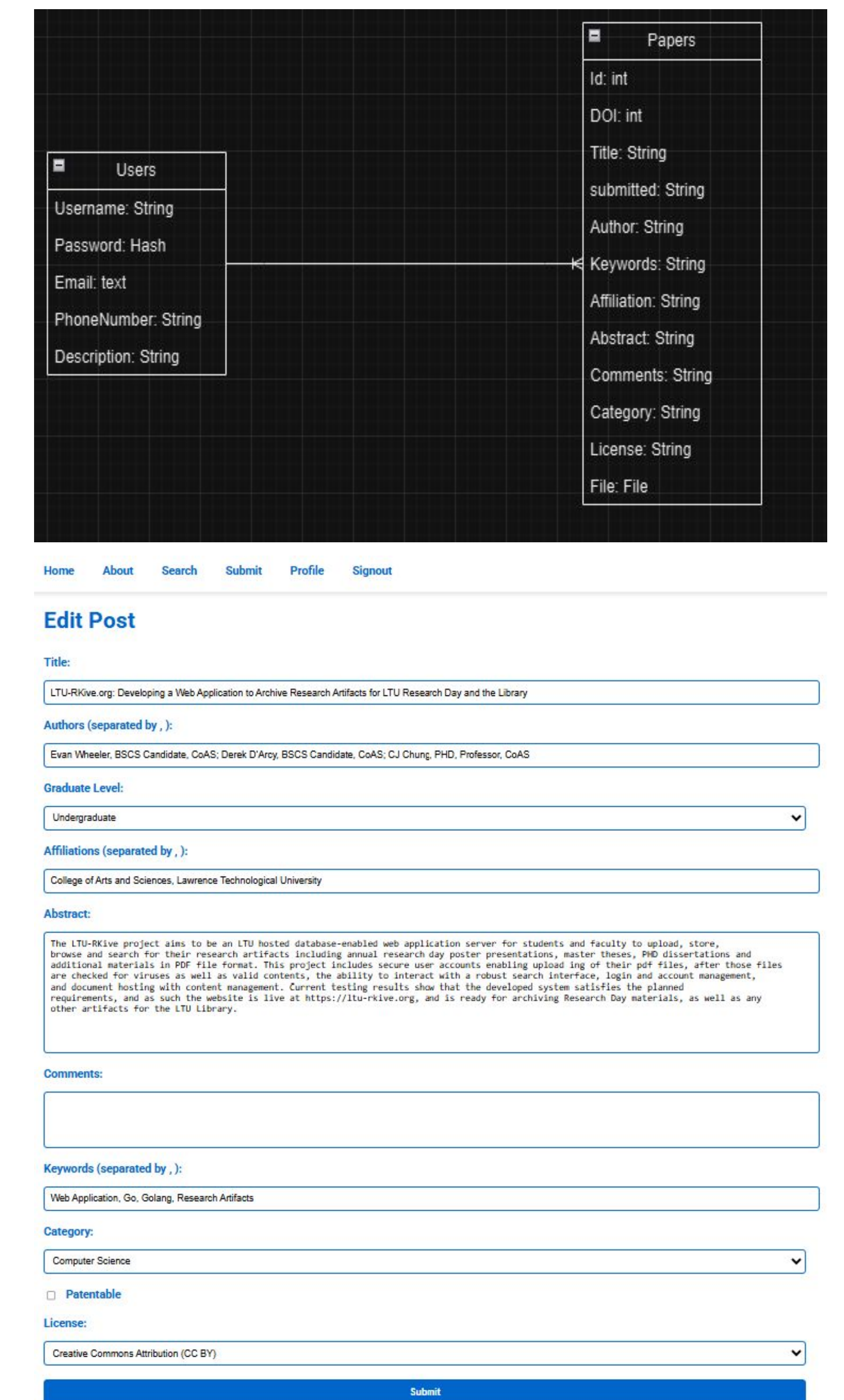

Conclusion

We ran into a few issues, such as the original LTU provided Debian server not being properly exposed due to firewall rules, which was corrected by switching to AWS. We also had some issues getting the virus checking API up and running, but once it got configured it works very smoothly to check if PDFs are infected. When setting up the email server, the process was not as smooth as we hoped. AWS blocks outgoing emails on port 25 by default to prevent spam, rendering the planned implementation unusable. While we await a response from Amazon, the current solution is to use a temporary email set up by Dr. Chung as a stop-gap measure, until a more permanent solution can be put in place.

Overall, this project was highly successful and current testing results show that the developed system at https://ltu-rkive.org is ready for archiving Research Day materials in PDF format.

References

[1] https://en.wikipedia.org/wiki/ArXiv
[2] https://go.dev/
[3] https://cloudmersive.com/
[4] https://www.doi.org/the-identifier/resources/faqs

Database Design

Future Work

There are a few solutions that could be implemented to take this project further. For the library to use the site, we need to update the system to handle other research artifacts besides Research Day presentations, such as MS theses and PhD Dissertations. The database would also need to be updated to add publication info such as publication date, venue, and publisher. The development of a solution to content check the uploaded PDFs using hand-crafted and Machine learning based algorithms was planned, but not implemented this semester. The current email sending function is temporary as well, and would need to be updated once a approval from AWS is received. We could also extend this project further by assigning DOI numbers[4] upon submission. This would take a joint effort with the school as an organization registering to be a "Registration Agency", or member of an RA in order to be able to supply DOI numbers. Another future task if and when this project is officialized would be to seamlessly integrate the site and service with the rest of the LTU campus website. This would allow the site to be truly connected with the rest of the campus services, and integrate the campus sign-on so that students and staff at LTU would not need a new account.