



IJIRCCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 9, Issue 5, May 2021

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 7.488

 9940 572 462

 6381 907 438

 ijircce@gmail.com

 www.ijircce.com

Mailbox Management to Auto Sort and Process Inbound Reports and Classify Documents

Harshith Kumar K, Dr. Jayasimha S R

PG Student, Dept. of MCA, RV College of Engineering, Bangalore, India

Assistant Professor, Dept. of MCA., RV College of Engineering, Bangalore, India

ABSTRACT: Automata is a system that was built to automate the process of manual traversal of our clients/ their vendors web portals where many of our workflow's data reports were available. Further to traversal of web pages, the system was designed to download and perform conversions of data and file types as per requirements of the process. The last activity the system performs is to upload the processed reports to the sFTP as instructed by the user from the team.

KEYWORDS: Automation, Vendors, AWS, Web Technology, Protocols.

I. INTRODUCTION

Clinical Trials generate lots of data - starting from the life cycle of the specimen till it is destroyed. A good amount of data is generated by different data points within the process. The problem grows many folds when there is not just one single source of data, but several – clients, and their vendors. Collection of this data becomes cumbersome and time consuming, and also no net value for manual effort when this has to be done by logging into the client's/ vendor's web portal, or sometimes even sourced through emails sent by them. At Global Specimen Solutions, a Covance division, a proprietary cloud-based specimen data management system called GlobalCode® supports the lifecycle management of the specimens by uniting multiple data sets like informed consent, specimen, clinical and biomarker data from many vendors for a client into one platform. Working with GSS, following will be the objectives of the intern. Learn and master Ruby, Ruby on Rails, and, Python (Selenium) to a working level. Understand the nuances of Rails' various components. Use the tool (Automata) that was developed in-house, as a base to develop automated browser actions. Build email-based automations which involve use of browser actions. Develop mailbox functionality on Automata so as to automate sorting, and processing of inbound emails from clients or vendors. Develop the base framework for classifying incoming (ICF) documents into 2 classes – Codifiable and non-Codifiable classes.

II. RELATED WORK

In this paper, the works involving the effort supported out by many authors work on various technology that are used to develop the web application Automata. The present paper discusses the need of automation testing in the process of software development, in order to provide high quality, robust and reliable software product. It answers the question why automation testing plays such a significant role in software development lifecycle as well as why not to use already existing automation testing tools when testing web applications and why it is better to create automation testing framework. Some reliable approaches how to build a testing framework are investigated. In recent years, there is advancement in the field of software engineering, applications are getting automated. As the software is error prone, there should some easy and automated way to test the software application. It is very challenging to test such complicated web applications. [3] Selenium is a web application testing tool and also it is an open-source freeware. This automation testing framework has gained wide acceptance as a popular and successful mode of website automated testing in a very short time span. These tools are widely used for testing graphical user interface and functionality of web-based applications developed for all types of industries such as e-commerce, travel, biotech, pharmaceuticals, and other mechanization. This testing freeware renders a cost-effective way which is an open-source testing framework for performance and other parameters to ascertain compatibility, accuracy, aspect and consumption of web applications. [4] Most of the time, customers request complex business logic to be implemented in software applications. Therefore, as long as business requirements grow, the pressure increases on the testing team to deliver the product with high quality

in a very tight time. Manual testing is not suitable for critical and complex applications in terms of both human resources and time. Therefore, there is a strong need to propose an automated testing framework which could reduce the overall software testing time. Automation testing has been introduced to overcome manual testing problems.[5] The advances in networking technologies and the increase in the need for storage resources have prompted many companies to outsource their storage needs. Cloud-storage providers offer clean and simple file-system interfaces, abstracting away the complexities of direct hardware management. At the same time, however, such services eliminate the direct oversight of performance that final users with high service-level requirements traditionally expect. While several works in literature have addressed security-related issues (such as privacy, integrity, availability, etc.) few of them have targeted the network performance of this kind of service.

III.SCOPE

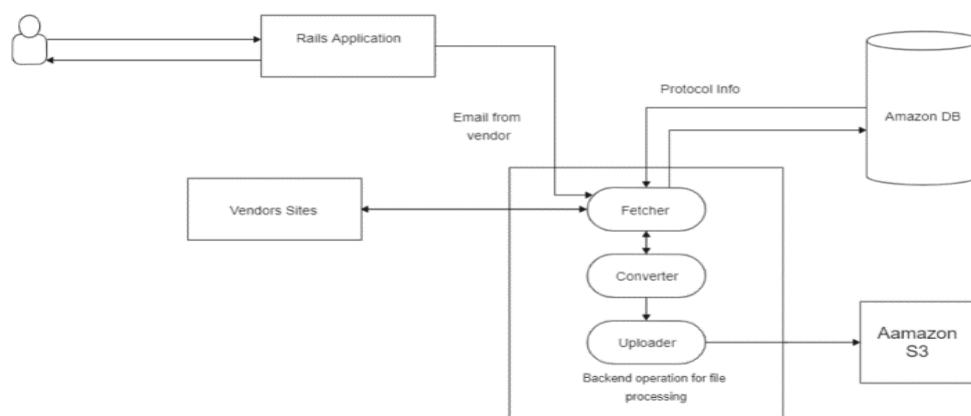
Automata this project is to automate the process of manual downloading of report from various vendor portals to process email which contains the reports to perform any file conversion and modification if necessary, to upload the result into secure file transfer protocol. The project is developed based on the principle of waterfall model the requirements are collected and made note in our internal project development tool and their progress will be tracked by means of this project development tool when all the requirement is completed, the release will be handed over to software testing and validation team for formal verification and validation when the formal verification and verification is completed the product will released to the end user.

IV.METHODOLOGY

Automata is a system that was built to automate the process of manual traversal of our clients/ their vendors web portals where many of our workflow’s data reports were available. Further to traversal of web pages, the system was designed to download and perform conversions of data and file types as per requirements of the process. The last activity the system performs is to upload the processed reports to the sFTP as instructed by the user from the team. All of this was put together by creating a web application built using Ruby on Rails, and processing parts taken care by Python. The web application was mainly used to provide an interface to store and edit from time to time – all the requirements in terms of downloading and converting of the various reports.

The web stack is straightforward with its sole functionality being a storehouse of data, and all of the automation was in actual executed by Python which involves three major functionalities performed across the various clients / their vendors and Protocols File Fetcher, File Manipulator/File Converter and File Uploader.

- **File Fetcher** mainly helps to fetch the file from the Vendor by collecting the Browser URL and all necessary navigation details in order to fetch the file from the vendor portal
- **File Manipulator** (File converter) mainly helps us to convert the given file format to CSV format with some data manipulation changes if required/specified by the client such as addition of time stamp at the end of the file name etc.
- **File Uploader** is used to upload the converted CSV format file to a Destination that is SFTP (Secure File Transfer Protocol) and also statuses of the file across each component are monitored easily through Automata application.



V. SIMULATION RESULTS

A. The Welcome Screen in Automata project where the user can see the various vendors and protocols list.

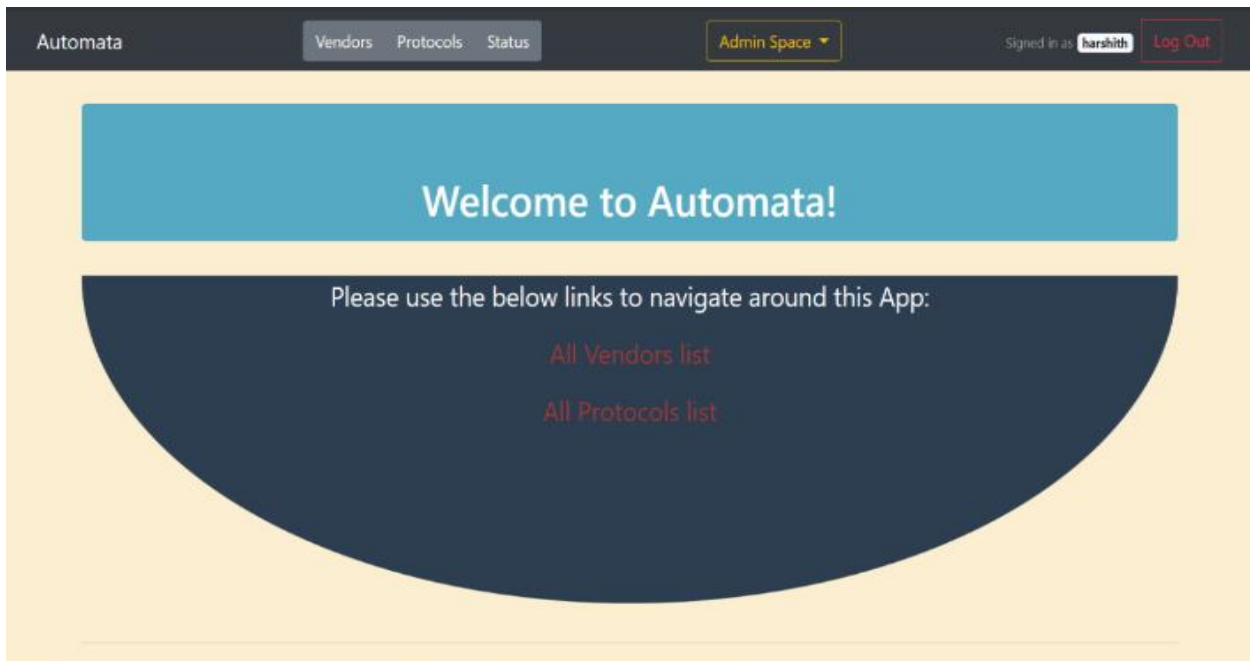


Fig 1.1 The Welcome screen

B. Selecting Various protocol and providing various required information.


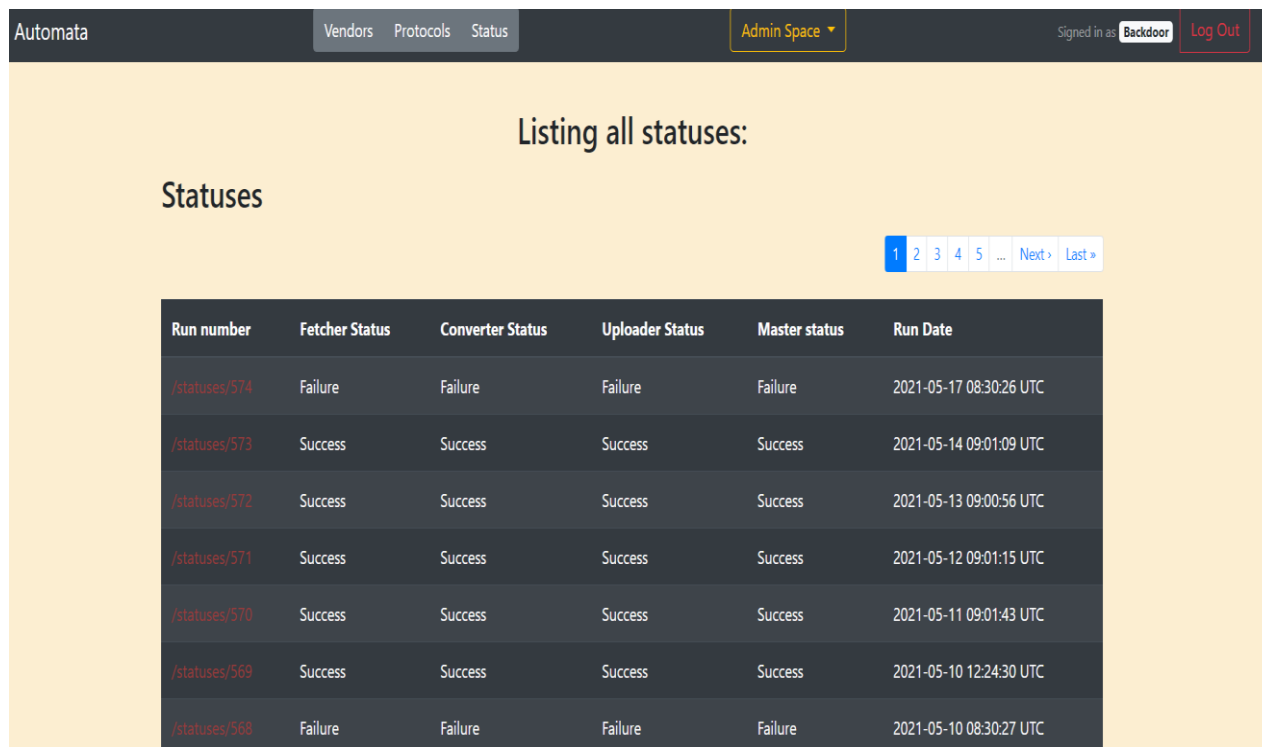


Fig 1.2 Protocol Screen

C. The Results of the various modules and its status for various vendor task.



The screenshot shows a web interface for 'Automata' with a navigation bar containing 'Vendors', 'Protocols', and 'Status'. The main content area is titled 'Listing all statuses:' and displays a table of status results. The table has columns for 'Run number', 'Fetcher Status', 'Converter Status', 'Uploader Status', 'Master status', and 'Run Date'. The data rows show a mix of 'Success' and 'Failure' statuses for various run numbers.

| Run number | Fetcher Status | Converter Status | Uploader Status | Master status | Run Date |
|---------------|----------------|------------------|-----------------|---------------|-------------------------|
| /statuses/574 | Failure | Failure | Failure | Failure | 2021-05-17 08:30:26 UTC |
| /statuses/573 | Success | Success | Success | Success | 2021-05-14 09:01:09 UTC |
| /statuses/572 | Success | Success | Success | Success | 2021-05-13 09:00:56 UTC |
| /statuses/571 | Success | Success | Success | Success | 2021-05-12 09:01:15 UTC |
| /statuses/570 | Success | Success | Success | Success | 2021-05-11 09:01:43 UTC |
| /statuses/569 | Success | Success | Success | Success | 2021-05-10 12:24:30 UTC |
| /statuses/568 | Failure | Failure | Failure | Failure | 2021-05-10 08:30:27 UTC |

Fig 1.3 The Various results of the modules

VI.CONCLUSION AND FUTURE WORK

The Automata or mailbox management to auto sort and process inbound reports and classify documents was developed for the sole purpose of eliminating the tedious process of manual work such as getting client data from various vendor or client website or from email etc. and converting it to the required format and uploading it to the S3 buckets in AWS. Automata eliminates all the manual work such as getting the client data converting it to required file type such as csv etc and upload it to the S3 bucket. The main goal or objective of the project is to eliminate manual work so that the various other process and can be completed swiftly.

REFERENCES

1. Zhang Qi, Cheng Lu and Boutaba Raouf, "Cloud Computing: State-of-the-art and research challenges", *J Internet ServAppl*, 2010.
2. Alberts, C., Dorofee, A., Stevens, J., & Woody, C. 2003. Introduction to the OCTAVE Approach. Carnegie Mellon University. Pittsburg, USA.
3. Alberts, D. S., & Garstka, J. 1999. Information Superiority and Network Centric Warfare. Retrieved April, 2005, from <http://www.iwar.org.uk/iwar/resources/info-superiority1999/index.htm>
4. Alberts, D. S., Garstka, J. J., Hayes, R. E., & Signori, D. A. 2001. Understanding Information Age Warfare: CCRP Publications, Washington, DC
5. Anderson, R. 2001. Why Information Security is Hard - An Economic Perspective. Paper presented at the Proceedings 17th Annual Computer Security Applications Conference. IEEE Computer Society, Los Alamitos, California.
6. Berger, B. 2003. Data-Centric Quantitative Computer Risk Assessment. Retrieved April, 2005, from <http://www.sans.org/rr/whitepapers/auditing/1209.php>
7. Fantuzzi, J. 2005. Document security? Tell me another joke. Retrieved 10/10/2005, 2005, from http://news.com.com/Document+security+Tell+me+another+joke/2010-1071_3-5783062.html?tag=st.ref.goo



8. Hutchinson, W., & Warren, M. 2001. Principles of Information Warfare. Journal of Information Warfare, Vol. 1, No. 1 pp 1-6.
9. Prasad Padhy Rabi, Patra Manas Ranjan and Chandra Satyapathy Suresh, "Cloud Computing: Security Issues and Research Challenges", *IJCSITS*, vol. 1, no. 2, December 2011.
10. Jensen Meiko and Sehwenk Jorg, "On Technical Security Issues in Cloud Computing", *IEEE International Conference on Cloud Computing*, pp. 109-116, October 2009.
11. Sengupta Shubhashis, Kaulgud Vikrant and Saujanya Vibhu, "Cloud Computing Security-Trends and Research Directions", *2011 IEEE World Congree on Services*.
12. Carroll Mariana, Merwe Alta van der and Kotze Paula, Secure Cloud Computing. Benefits Risks and Controls, IEEE, 2011.
13. Chen Deyan and Zhao Hong, *Data Security and Privacy Protection Issues in Cloud Computing. 2012 International Conference on Computer Science and Electronics Engineering*, 2012.
14. Behl Akhil and Behl Kanika, *An analysis of Cloud Computing Security Issues. 2012 IEEE*.
15. Khan Amreen and Ahirwar Kamal Kant, "Mobile Cloud Computing as a Future of Mobile Multimedia Database", *IJCSC*, vol. 2, no. 1, pp. 219-221, January-June 2011.
16. Lombardi Flavio and Di Piero Roberto, "Journal of Network and Computer Applications", *Journal of Network and Computer Applications*.
17. Qian Ling, Luo Zhiguo, Du Yujian and Guo Leitao, *Cloud Computing: An Overview*.
18. Mladen A. Vouk, "Cloud Computing- Issues Research and Implementations", *Journal of Computing and Information Technology -CIT 16*, vol. 4, pp. 235-246, 2008.
19. Khurana Sumit and Gaurav Verma Anmol, "Comparison of Cloud Computing Service Models: SaaS PaaS IaaS", *IJECT*, vol. 4, April-June 2013.



INNO SPACE
SJIF Scientific Journal Impact Factor

Impact Factor:
7.488

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

 9940 572 462  6381 907 438  ijircce@gmail.com



www.ijircce.com

Scan to save the contact details