A
Project Report
on

"Effective End-of-Life Management of Assets in Organization by Snowflake Integration"

Submitted to

Sant Gadge Baba Amravati University, Amravati

Submitted in partial fulfilment of
the requirements for the Degree of
Bachelor of Engineering
in
Computer Science & Engineering

Submitted by

Ms. Surabhi Lahoti
(PRN: 203120182)

Mr. Kunal Chandore
(PRN: 203120279)

Mr. Aditya More
(PRN: 203120330)

Under the Guidance of
Dr. J. M. Patil
Assoc. Prof., CSE Dept.

Department of Computer Science and Engineering
Shri Sant Gajanan Maharaj College of Engineering,
Shegaon – 444 203 (M.S.)

Session 2023-2024

A
Project Report
on

"Effective End-of-Life Management of Assets in Organization by Snowflake Integration"

Submitted to

Sant Gadge Baba Amravati University, Amravati

Submitted in partial fulfilment of
the requirements for the Degree of
Bachelor of Engineering
in
Computer Science & Engineering

Submitted by

Ms. Surabhi Lahoti
(PRN: 203120182)

Mr. Kunal Chandore
(PRN: 203120279)

Mr. Aditya More
(PRN: 203120330)

Under the Guidance of
Dr. J. M. Patil
Assoc. Prof., CSE Dept.

Department of Computer Science and Engineering
Shri Sant Gajanan Maharaj College of Engineering,
Shegaon – 444 203 (M.S.)

Session 2023-2024

SHRI SANT GAJANAN MAHARAJ COLLEGE OF ENGINEERING,

SHEGAON – 444 203 (M.S.)

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

EST.1983

CERTIFICATE

This is to certify that **Ms. Surabhi Lahoti, Mr. Kunal Chandore** and **Mr. Aditya More** student of final year Bachelor of Engineering in the academic year **2023-24** of Computer Science and Engineering Department of this institute have completed the project work entitled **"Effective End-of-Life Management of Assets in Organization by Snowflake Integration"** based on syllabus and has submitted a satisfactory account of their work in this report. Hence recommended for the partial fulfillment of degree of Bachelor of Engineering in Computer Science and Engineering.

Dr. J. M. Patil
Project Guide

Ms. N. N. Ghuikar
Co-Guide

Dr. J. M. Patil
Head of Department
CSE

Dr. S. B. Somani
Principal
SSGMCE, Shegaon

SHRI SANT GAJANAN MAHARAJ COLLEGE OF ENGINEERING,

SHEGAON – 444 203 (M.S.)

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

CERTIFICATE

This is to certify that **Ms. Surabhi Lahoti, Mr. Kunal Chandore** and **Mr. Aditya More** student of final year Bachelor of Engineering in the academic year **2023-24** of Computer Science and Engineering Department of this institute have completed the project work entitled **"Effective End-of-Life Management of Assets in Organization by Snowflake Integration"** based on syllabus and has submitted a satisfactory account of their work in this report. Hence recommended for the partial fulfillment of degree of Bachelor of Engineering in Computer Science and Engineering.

Internal Examiner

Dr. J. M. Patil

Name and Signature

Date: 10/05/24

External Examiner

Dr. G. J. Sawale

Name and Signature

Date: 10/05/2024

Acknowledgement

It is our utmost duty and desire to express gratitude to various people who have rendered valuable guidance during our project work. We would have never succeeded in completing our task without the cooperation, encouragement and help provided to us by then. Several people deserve recognition for their unwavering support and guidance throughout this report.

We are highly indebted to our guide **Dr. J. M. Patil** for his guidance and constant supervision as well as for providing necessary information from time to time. We would like to take this opportunity to express our sincere thanks, for his esteemed guidance and encouragement. His suggestions broadened our vision and guided us to succeed in this work.

We are sincerely thankful to **Dr. J. M. Patil** (HOD, CSE Department, SSGMCE, Shegaon), and to **Dr. S. B. Somani** (Principal, SSGMCE, Shegaon) who always has been kind to extend their support and help whenever needed.

We would like to thank all teaching and non-teaching staff of the department for their cooperation and help. Our deepest thanks to our parents and friends who have consistently assisted us towards the successful completion of our work.

Projectees

Ms. Surabhi Lahoti

Mr. Kunal Chandore

Mr. Aditya More

Final Year B.E. Sem-VIII

Session 2023-24

Abstract

To keep devices stable and secure in the modern industrial environment, vulnerabilities must be found and fixed. One of the biggest challenges in asset management is keeping track of the constantly changing End-Of-Life (EOL) statuses and navigating through many data sources. The lack of a single dashboard makes it more difficult to obtain a comprehensive overview of all relevant data. We suggest putting in place a scraping method to collect data from various sources in order to overcome these issues. A consolidated, scalable, and dependable dashboard that easily connects into current systems will show the outcomes. One of the main reasons for creating a centralized comprehensive dashboard is the desire to improve and simplify a highly fragmented and disjointed process, as well as the requirement for clarity and simplification in the management of process information. Advanced data slicing and dicing techniques will be used, utilizing Snowflake's capabilities, to handle enormous datasets efficiently.

This all-encompassing strategy seeks to improve asset management procedures, strengthen device security, and offer useful information for defensible decision-making. Organizations can simplify difficult tasks and guarantee device stability and security by efficiently controlling EOL statuses and combining data via a consolidated dashboard.

Keywords – End-of-Life (EOL), Centralized Dashboard, Data Slicing and Dicing, Asset Management, Snowflake

Contents

Title	Page No.
Abstract	*i*
Contents	*ii*
List of Figures	*iii*
List of Abbreviations	*iv*
Chapter -1: Introduction	**01**
1.1 Preface	1
1.2 Background of the Study	1
1.3 Problem Statement	2
1.4 Objectives	2
1.5 Scope of the project	3
1.6 Organization of the Project	3
Chapter -2: Literature Survey	**06**
Chapter -3: Proposed Methodology	**09**
3.1 Web Scraping Mechanism	9
3.2 Data Storage and Management	12
3.3 Snowflake Integration	13
3.4 Scalability and Performance Optimization	15
3.5 Privacy and Data Security	16
3.6 Centralized Dashboard	16
Chapter -4: Results and Discussion	**18**
Chapter -5: Conclusion	**19**
Chapter -6: Future Scope	**20**
References	**21**
Dissemination of Work	**24**
Publication Details	24
Research Paper	27
Sponsorship Details	35

Plagiarism Report (using Turnitin software) **38**

List of Figures

Figure No.	Name of the Figures	Page No.
Figure 1	The Approach of our Study.	9
Figure 2	Web Scraping Code using Selenium.	11
Figure 3	Main Logic behind Web Scraping.	11
Figure 4	Folder Structure.	13
Figure 5	Snowflake Architecture.	14
Figure 6	The Snowflake Scoreboard for Total Devices Count.	18
Figure 7	The Snowflake Graph for Devices according to Vendor.	18

List of Abbreviations

Abbreviation	Description	Page No.
MDM	Master Data Management	01
ITAM	IT Asset Management	01
EOL	End-of-Life	01
WS	Web Scraping	06
HTML	Hyper Text Markup Language	06
FQDN	Fully Qualified Domain Name	08
CIDR	Classless Inter-Domain Routing	08
HTTP	Hypertext Transfer Protocol	10
SQL	Structured Query Language	12
CSV	Comma-Separated Values	12
ETL	Extract, Transform, and Load	13
RBAC	Role-Based Access Control	16
PII	Personally Identifiable Information	16
GDPR	General Data Protection Regulation	16
HIPAA	Health Insurance Portability and Accountability Act	16
PCI	Payment Card Industry	16
DSS	Data Security Standard	16
UI	User Interface	17

Chapter 1

INTRODUCTION

1. INTRODUCTION

1.1 Preface

Ensuring compliance and receiving real-time updates on device status are pivotal aspects of maintaining organizational security, especially in light of recent findings from our company's report. The report indicates that a significant portion of security teams are tasked with overseeing an extensive number of devices, with a quarter managing upwards of 250,000 devices and over 10,000 workstations. However, despite the critical nature of device management, only 23% of security teams have fully enrolled their organization's devices in their Master Data Management (MDM) solution. Additionally, merely 25% have ensured that all or nearly all of their devices are running the latest operating system.

To address these issues, a proposed solution involves the implementation of a scraping mechanism designed to collect data from multiple sources [10][15]. Following rigorous verification and EOL updates, the outcomes are presented through a centralized, scalable, and reliable dashboard in Snowflake [9]. This dashboard, designed for seamless integration, becomes the cornerstone for streamlined asset management.

1.2 Background of the Study

User The study originates from the imperative to address the multifaceted challenges associated with digital transformation in asset management, especially within the domain of IT Asset Management (ITAM). With the proliferation of digital assets and the complexities surrounding End-of-Life (EOL) asset management, organizations are grappling with the need for efficient and effective solutions. Leveraging innovative technology, the study aims to automate the extraction and management of EOL asset data by utilizing advanced techniques such as web scraping and Snowflake's data warehousing solution [9][10][11]. This approach is motivated by the recognition that manual processes are no longer sufficient to handle the vast volumes of data generated by modern asset management practices.

In a comprehensive exploration of the digital transformation landscape in asset management, the study adopts a qualitative Delphi methodology involving 15 experts drawn from academia, consultancy, and industry. This diverse panel of experts contributes valuable insights and perspectives, facilitating the identification,

validation, and categorization of the drivers and obstacles influencing digital transformation endeavours [12]. Through collaborative interactions with experts, the study pinpoints critical barriers impeding digital transformation, including strategic misconceptions, vision deficiencies, entrenched cultural norms, and skills gaps. Moreover, the study illuminates pivotal drivers such as cost efficiencies, opportunities in asset monitoring, and anticipated benefits in risk management processes.

The anticipated outcomes of this study hold the promise of significantly enriching the existing body of knowledge concerning digital transformation in asset management. By uncovering the factors shaping digital transformation initiatives, the study aspires to equip organizations with actionable insights to navigate their digital transformation journey with greater efficacy. Ultimately, the research aims to elevate operational efficiency, refine asset management methodologies, and fortify the security of EOL assets, empowering organizations to thrive amidst the ever-evolving digital landscape.

1.3 Problem Statement

In contemporary industries, ensuring device security and system stability is imperative, requiring the prompt identification and resolution of vulnerabilities. However, managing the fluctuating End-Of-Life (EOL) statuses and numerous data sources complicates asset management processes. Furthermore, the absence of a centralized dashboard obstructs the comprehensive visibility of critical data. To tackle these obstacles, our proposed solution involves implementing a scraping mechanism to gather data from various sources. Subsequently, following the verification and updating of EoL statuses, the results will be presented in a centralized, scalable, and dependable dashboard that seamlessly integrates with existing systems. To efficiently handle the vast dataset, advanced data slicing and dicing techniques using Snowflake will be employed.

1.4 Objectives

1. To implement robust vulnerability management mechanisms that enable the proactive identification and mitigation of potential security threats. This will enhance the overall security posture of the system.

2. To develop efficient methods for monitoring and managing the changing End-of-Life (EOL) status of devices. This will ensure timely replacements and upgrades, minimizing disruptions and vulnerabilities.

3. To establish a standardized approach for handling massive and diverse datasets. This involves implementing procedures for data cleansing, integration, and storage, ensuring data quality and accessibility.

4. To design and build a dynamic and scalable dashboard capable of effectively managing the expanding dataset. This dashboard will provide real-time insights and visualizations, enabling informed decision-making.

5. To create an intuitive and user-friendly dashboard interface. This interface will feature interactive visualizations and customized filters, enhancing usability and making data interpretation more accessible.

By achieving these objectives, we aim to enhance security, streamline device management, optimize data handling, and provide a user-friendly experience through our project. These goals are measurable, time-bound, and aligned with the overarching mission of the project.

1.5 Scope of the project

1. A standardized approach will be established for handling massive and diverse datasets.
2. Design and build a dynamic and scalable dashboard capable of effectively managing the expanding dataset.
3. Identifying robust mechanisms for identifying and mitigating potential security threats proactively.

1.6 Organization of the project

Chapter 1: Introduction: It gives an introduction of the project. The introduction provides a comprehensive overview of the project's objectives, methodology, and scope, laying the groundwork for addressing critical challenges in organizational security and device management.

Chapter 2: Literature Survey: Literature review of the research papers referred to get an idea of the previous work. The survey also references significant research efforts in cybersecurity, particularly regarding the security of EOL embedded devices. It emphasizes the importance of continued research in addressing evolving threats associated with EOL devices.

The literature survey provides a comprehensive summary of existing knowledge, highlighting important themes, arguments, and gaps in the literature relevant to the research project's objectives.

Chapter 3: Methodology: Refers to the systematic, structured approach used to plan, execute, and manage a project from start to finish. It encompasses the processes, techniques, tools, and guidelines employed to achieve the project's objectives efficiently and effectively. Methodology provides a framework for organizing tasks, managing resources, mitigating risks, and ensuring quality throughout the project lifecycle.

Chapter 4: The Results and Discussion: This section presents the findings of the study and offers analysis and interpretation of those results. The results and discussion section serves as the heart of a research study, presenting the empirical findings, analysing their significance, and offering insights into their broader implications. It represents the culmination of the research process and provides a foundation for advancing knowledge and informing decision-making in relevant domains.

Chapter 5: Conclusion: The conclusion drawn from the "Effective End-of-Life (EOL) Management in Organization" project underscores its transformative impact on the organization, marking a pivotal shift towards proactive vulnerability management and dynamic EOL status management. Through the implementation of advanced strategies like the scraping mechanism and the development of a centralized dashboard, tangible outcomes have been achieved, empowering stakeholders with unified, data-driven insights.

Chapter 6: Future Scope: Future scope refers to the potential opportunities and prospects in a particular field, industry, or career. It's essential to consider because it helps individuals make informed decisions about their education, career, and personal development.

To make the most of future scope assessments, use them as a foundation for setting goals, making informed decisions, and crafting a flexible career strategy. Continuously update your knowledge and skills to remain competitive in your chosen field.

References: Research papers referenced in a project play a crucial role in shaping its conceptual framework, informing its methodology, guiding data analysis, validating findings, supporting discussions, and upholding academic integrity. In a project like this, referencing research papers serves several purposes, including providing evidence for the proposed solution, supporting the background and problem statement, and informing the methodology and approach. Here is an elaboration on how research papers are referenced and used in the context of this project.

Chapter 2

LITERATURE SURVEY

2. LITERATURE SURVEY

This part lays out the theoretical foundation for the research endeavour, gives a thorough summary of the state of knowledge at the moment, and highlights important themes, arguments, and gaps in the literature. Ensuring compliance and real-time device status updates are critical for organizational security, particularly given recent findings from our company's report. It reveals that a significant portion of security teams oversee extensive numbers of devices, with 25% managing over 250,000 devices and 10,000 workstations. However, only 23% have fully enrolled devices in their Master Data Management (MDM) solution, and merely 25% ensure all devices run the latest operating system [8]. These statistics highlight challenges posed by current MDM solutions, lacking comprehensive functionality, hindering full device visibility, compliance, and real-time monitoring. Existing solutions often provide incomplete data, impeding understanding of an organization's device landscape, while the absence of feedback mechanisms complicates device management.

In the digital era, marked by the information age, there's been a widespread shift to digital technologies, leading to the rise of intangible artifacts integrated into daily life. Despite the immunity of intangible systems to physical degradation, software still has a finite lifespan, posing challenges in identifying its End-of-Life (EOL). However, current literature lacks sufficient discussion on methods for determining software EOL and lacks exploration of decommissioning methods.[11]

The internet is a vast information hub, accommodating diverse needs from business to personal use. Web Scraping (WS) plays a crucial role in extracting large amounts of data from websites, transforming unstructured HTML into structured data. This paper explores various web scraping techniques, including Selenium, requests Library, and third-party libraries, shedding light on their efficacy and ethical considerations. Categorized results reveal website vulnerability to bot attacks and differing security levels. To improve accuracy, the paper suggests utilizing additional modules and conducting tests on a larger scale, while advocating for further assessment of Selenium-based tests with additional parameters. [10]

Recent studies have expanded the scope of Selenium Automation beyond software testing, emphasizing its effectiveness in web scraping and data extraction. Integrating techniques like Count Vectorizer Machine Learning and Forward Selection Wrapper methods with Selenium enhances its ability to retrieve relevant information from web pages. Despite its versatility, challenges such as data extraction inaccuracies and the requirement for manual intervention persist. Nonetheless, Selenium's broader application in streamlining data retrieval tasks highlights its potential in domains like marketing and client services.[15]

In today's business landscape, data holds unparalleled importance, driving decision-making across industries. To effectively harness the vast volumes of data generated, businesses require tools that offer high performance, scalability, and ease of use. Snowflake Elastic Data Warehouse emerges as a leading solution, providing organizations with the capabilities to manage large datasets efficiently. Through its architecture, data sharing, storage, ingestion, transformation, analytics, and visualization features, Snowflake stands out as the preferred choice for handling extensive data volumes with exceptional speed and performance.[9]

The study carried out by Jiang et al. and Wang et al. [14] is a significant attempt to look into the (in)security of End-of- Life embedded devices. The aliveness analysis and vulnerability analysis provide a holistic view of the challenges posed by active EoL devices, contributing to the body of knowledge in the field of cyber security. This literature review highlights the importance of their work, emphasizing the need for continued research and collaborative efforts to address the evolving threats associated with EoL devices.

The paper aims to explore the factors influencing digital transformation in asset management through a qualitative Delphi study involving 15 experts. It identifies 20 barriers, including strategic misunderstandings and inadequate asset management systems, along with 12 critical drivers such as cost reduction and opportunities in condition monitoring. These findings provide valuable insights for organizations undertaking digital transformation initiatives in asset management, highlighting the importance of addressing both barriers and drivers effectively.[5]

Asset owner identification is paramount for information security organizations, aiding in breach detection, vulnerability assessment, and countermeasure definition. This research employs various machine learning algorithms to predict asset ownership, conducting separate analyses for each owner and utilizing a 100- iteration Monte Carlo Cross Validation. The resulting visualization dashboard enables exploratory data analysis and model evaluation. Ad boost emerges as the top performing model, while Naïve Bayes lags. Key features like Fully Qualified Domain Name (FQDN), CIDR, and location significantly impact ownership prediction.[3]

The evolution of technology has revolutionized traditional business practices, particularly with the emergence of e-commerce websites as prominent platforms for conducting transactions. This transition has significantly impacted the dynamics among marketers, consumers, and retailers, collectively referred to as consumers in this context, by providing convenient avenues for buying and selling goods online. The proliferation of e-commerce websites has simplified product search operations, facilitating tasks such as accessing updated information on new or edited products, sales, and costs. Without employing methods like the one presented in this paper or those discussed in existing literature, navigating through e-commerce marketplaces for activities like price comparison would be time-consuming. Given the rapidly changing nature of prices and product information, it is imperative to analyze competitor prices and intermediary sellers' roles regularly to make informed decisions. [13]

Chapter 3

PROPOSED METHODOLOGY

3. PROPOSED METHODOLOGY

The first step involves identifying a comprehensive list of End-of-Life (EoL) models from various vendors. This selection process includes prominent vendors such as Lenovo, Dell, Hitachi among others. EoL models are chosen based on their prevalence and significance in the market, ensuring a representative sample for analysis. Web scraping is employed to extract relevant data regarding EoL dates and device information from the official websites of selected vendors. Selenium, a powerful automation tool, is utilized for web scraping due to its capability to interact with dynamic web elements. Selenium scripts are developed to navigate through vendor websites, locate pertinent data, and extract it systematically.

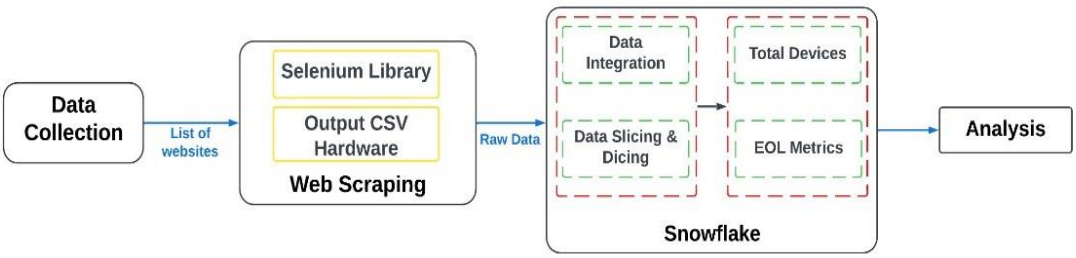

Figure 1: The Approach of our Study.

3.1. Web Scraping Mechanism:

The process for extracting data from web pages is known as web scraping. The methodology involves:

3.1.1. Data Collection:

Identifying the target websites and defining the data to be extracted. This includes understanding the website structure and elements. This involves identifying sources that contain relevant information for their study objectives. Furthermore, researchers need to define the specific data points they intend to collect, ensuring clarity on the information needed. Understanding the structure of the target websites is crucial at this stage, as it enables researchers to navigate the site effectively and locate the desired data elements. By comprehensively assessing the website's layout and organization, researchers can develop a strategy for efficient data extraction while ensuring the coverage of all pertinent information.

3.1.2. Use of Libraries:

Utilizing tools like Selenium, Beautiful Soup, or Scrapy in Python for efficient data extraction. Selenium is often used for dynamic content loading through automation. Once the target websites and data elements are identified, researchers can leverage a variety of web scraping libraries and tools to facilitate the data extraction process. Tools such as Selenium, Beautiful Soup, and Scrapy offer distinct capabilities for extracting data from web pages. Selenium, for instance, is particularly useful for handling websites with dynamic content or interactive features, as it enables automated interactions with the browser. Researchers can employ Selenium scripts to navigate through web pages, interact with form fields, and extract data from dynamically generated elements. Similarly, Beautiful Soup and Scrapy provide powerful parsing capabilities, allowing researchers to extract data from HTML documents with ease.

3.1.3. Programming Languages:

Python is a commonly used language due to its rich ecosystem of web scraping libraries and tools. It allows for handling HTTP requests, parsing HTML, and automating browser interactions. Python emerges as a prevalent choice for implementing web scraping mechanisms due to its versatility and extensive support for web scraping libraries and tools. Python offers a rich ecosystem of libraries for handling HTTP requests, parsing HTML documents, and automating browser interactions, making it well-suited for web scraping tasks. Researchers can utilize Python scripts to send HTTP requests to target websites, retrieve HTML content, and extract relevant data using libraries such as Requests and Beautiful Soup. Additionally, Python's support for automation frameworks like Selenium enables researchers to automate browser interactions, streamlining the data extraction process further. Overall, Python provides a robust and flexible platform for developing web scraping solutions that meet researchers' specific requirements and objectives.

This image showcases the Python code snippets used for web scraping EOL dates from manufacturer websites. The code utilizes the Selenium library for automated navigation and data extraction, demonstrating the process of retrieving and storing EOL information in the Snowflake database for further analysis and management.

```
"""Dell"""

from selenium import webdriver
from selenium.webdriver.common.by import By
import pandas as pd
from datetime import datetime
import time

options = webdriver.ChromeOptions()
options.add_argument("--headless")
options.add_argument("--window-size=1920,1080")


class Dell_Scraper:
    """Dell Scraper Class"""

    url = "https://relutech.com/eol-eosl/dell/?orderby=date-asc"

    def __init__(self):
        self.driver = None

    def open_url(self, url):
        """Common Function for opening the URL in Chrome"""
        self.driver = webdriver.Chrome("Program Files\chromedriver.exe")
        self.driver.get(url)
        time.sleep(10)
```

Figure 2: Web Scraping Code using Selenium.

```
    def eol_data_generator(self, url, output_filename):
        """scrape_data function collects the data
        : returns dataframe and CSV file"""
        self.open_url(url)

        table = self.driver.find_element(By.XPATH, "//table")
        print(table.text)
        mainList=[]
        rows=table.find_elements(By.XPATH, './/tr')
        for row in rows:
            tempList=[]
            data=row.find_elements(By.XPATH, ".//td | .//th")
            for datum in data:

                tempList.append(datum.text)

            mainList.append(tempList)

        df = pd.DataFrame(mainList[1:], columns=mainList[0])

        df["EOSL Date"] = df["EOSL Date"].apply(lambda x: self.convert_and_fill_date(x))
        # df.drop(df[df['End-of-Life'] == 'See note'].index, inplace=True)
        df['Source URL'] = url
        df['Vendor'] = 'Dell'

        #Converted dataframe to CSV
        df.to_csv(output_filename, index=False)
        print(df)
        self.driver.close()
        return df


#Calling the function and getting the CSV file
scraper = Dell_Scraper()
scraper.eol_data_generator(scraper.url, "CSV\HW\Dell.csv")
```

Figure 3: Main Logic behind Web Scraping.

3.2. Data Storage and Management:

After scraping, the data needs to be stored and managed effectively. The methodology includes:

3.2.1 Database or Data Warehouse:

Choosing an appropriate database or data warehouse for storage. Options consist of SQL databases (like PostgreSQL or MySQL), NoSQL databases (like MongoDB), or a cloud-oriented data warehouses (like Amazon Redshift, Google Big Query). Upon extracting data from web pages, researchers must deliberate on the optimal storage solution to accommodate the harvested information. In our project, Snowflake serves as the chosen database platform due to its unparalleled scalability, performance, and flexibility in handling diverse data types. Unlike traditional SQL databases, Snowflake offers a cloud-native architecture, enabling seamless scaling of storage and computational resources based on demand. Its support for structured, semi-structured, and unstructured data makes it an ideal choice for storing the varied data formats encountered during web scraping. Moreover, Snowflake's data sharing capabilities facilitate collaboration and data exchange across teams and organizations, enhancing the project's interoperability and data accessibility.

3.2.2 Data Cleansing:

Ensuring data quality by identifying and correcting errors or inconsistencies. This entails standardizing formats, eliminating duplicates, and handling missing values. Data cleansing is a critical step in the data management process aimed at ensuring the quality and reliability of the stored data. During this phase, researchers systematically identify and rectify errors, inconsistencies, and inaccuracies present in the scraped data. This may involve standardizing data formats to ensure uniformity, removing duplicate records to prevent redundancy, and addressing missing or erroneous values through imputation or deletion.

This figure displays the hierarchical folder structure used for organizing the project files related to the web scraping. The structure includes directories for web scraping program files and output CSV files.

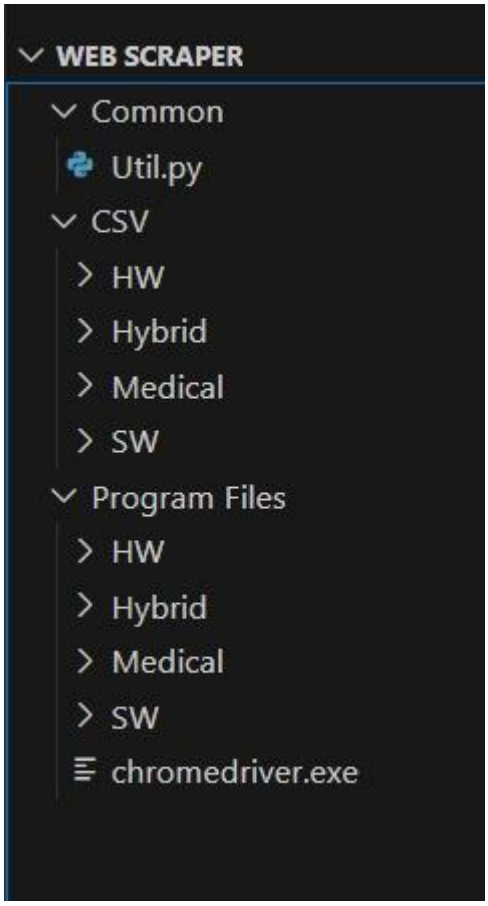

Figure 4: Folder Structure.

3.3. Snowflake Integration:

Integrating with Snowflake, a cloud-based data platform, involves several steps:

3.3.1. Data Loading:

Loading the scraped and cleansed data into Snowflake. This may include using Snowflake connectors or ETL (Extract, Transform, Load) tools. The initial step in integrating with Snowflake entails loading the scraped and cleansed data into the platform. This process can be facilitated using various methods, including Snowflake connectors that enable seamless data transfer from external sources to Snowflake. Alternatively, researchers may opt to utilize ETL (Extract, Transform, Load) tools compatible with Snowflake to automate the data loading process. By leveraging these tools and connectors, researchers can efficiently ingest large volumes of data into Snowflake, ensuring its availability for subsequent processing and analysis.

3..3.2. Data Processing:

Performing necessary transformations on the data within Snowflake to prepare it for analysis. This could involve aggregations, joins, or other operations. Once the data is loaded into Snowflake, we may perform necessary transformations to prepare it for analysis. This involves executing a series of data processing tasks within Snowflake, such as aggregations, joins, or data cleansing operations. By leveraging Snowflake's powerful SQL-based querying capabilities, researchers can manipulate the data to derive meaningful insights and patterns. Moreover, Snowflake's ability to handle complex data processing tasks at scale ensures that researchers can efficiently process large datasets without compromising performance or scalability.

3.3.3. Data Transformation:

Adapting the data to meet specific reporting or analysis requirements. This may involve formatting the data into structured formats, creating derived metrics or dimensions, or enriching the dataset with additional contextual information. By tailoring the data to align with the project's objectives and analytical needs, researchers can enhance the relevance and usability of the dataset for subsequent analysis and visualization tasks.

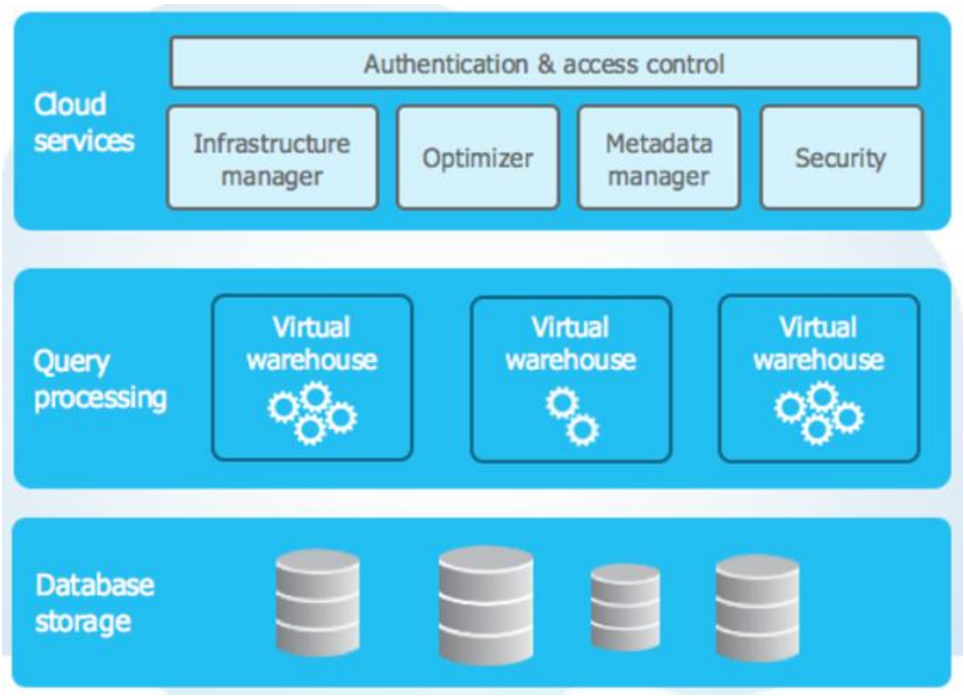

Figure 5: Snowflake Architecture.

3.4. Scalability and Performance Optimization:

Ensuring the system can handle growth and optimizing performance involves:

3.4.1. Database Scaling:

Scaling the database infrastructure to handle increased data volumes. This may involve horizontal scaling (adding more servers) or vertical scaling (upgrading hardware). As the volume of data processed and stored within the system grows, it becomes imperative to scale the underlying database infrastructure to accommodate this growth effectively. Database scaling can be achieved through horizontal scaling, which involves adding more servers or instances to distribute the workload across multiple nodes.

3.4.2. Query Optimization:

Optimizing database queries to reduce response times. This includes indexing, query caching, and other performance-tuning techniques. Optimizing database queries plays a crucial role in enhancing system performance and reducing response times. This involves employing various techniques such as indexing, which accelerates data retrieval by creating efficient data access paths. Additionally, query caching can be utilized to store the results of frequently executed queries, reducing the need for redundant computations and improving overall query performance. By implementing query optimization strategies, researchers can streamline database operations and enhance the efficiency of data retrieval and processing tasks.

3.4.3. Dashboard Performance:

Optimizing the performance of visualization tools and dashboards is essential for ensuring a responsive and user-friendly experience for stakeholders accessing analytical insights. This may involve implementing caching mechanisms to store precomputed aggregations or frequently accessed data, reducing the time required to render visualizations. By prioritizing dashboard performance optimization, researchers can facilitate seamless data exploration and analysis, empowering stakeholders to derive valuable insights efficiently.

3.5. Privacy and Data Security:

Maintaining the privacy and security of the data involves:

Access Control: Role-based access control (RBAC) can be used to limit access to sensitive data according to user roles. Role-based access control (RBAC) is a method used to restrict access to sensitive data based on predefined user roles within an organization. With RBAC, access permissions are assigned to specific roles, and users are then assigned to one or more roles based on their responsibilities and requirements. For example, administrative staff may have access to all data and system functionalities, while regular employees may only have access to data relevant to their specific job roles. RBAC helps enforce the principle of least privilege, ensuring that users only have access to the data and resources necessary to perform their duties, thereby reducing the risk of unauthorized access and data breaches.

Compliance: Adhering to data protection regulations and industry-specific compliance standards. This may involve encryption, anonymization, and audit trails. Adhering to data protection regulations and industry-specific compliance standards is essential for organizations to ensure the security and privacy of sensitive information. Compliance measures may include implementing encryption techniques to protect data both at rest and in transit, anonymizing personally identifiable information (PII) to prevent unauthorized disclosure, and maintaining audit trails to track data access and modifications. By following compliance standards such as GDPR, HIPAA, or PCI DSS, organizations can demonstrate their commitment to protecting customer data and mitigating the risk of regulatory penalties or legal liabilities.

3.6. Centralized Dashboard:

Creating a centralized dashboard involves designing an intuitive and user-friendly interface (UI) that provides stakeholders with access to key insights and analytics derived from the collected data. A well-designed UI enhances user experience by enabling easy navigation and interaction with the dashboard's features. Visualization plays a crucial role in presenting complex data in a visually appealing and comprehensible manner. By using charts, bar plots, scoreboard and other visualization techniques, organizations can effectively communicate trends, patterns, and correlations within the data, facilitating better decision-making.

User Interface (UI): Designing an intuitive and user- friendly interface for the dashboard.

Visualization: Choosing appropriate data visualization techniques to effectively communicate insights.

Filters and Interactivity: Implementing filters and interactive components that let users dynamically explore and analyze the data.

Chapter 4

RESULT AND DISCUSSION

4. RESULT AND DISCUSSION

This section presents visual representations of the results obtained from the tests conducted. These graphs serve to enhance the comprehension of the article's objectives by providing clear and accessible insights into the data analysis performed.

The seamless integration of Selenium web scraping and Snowflake data management effectively pinpointed more than 2,000 active End-of-Life (EoL) devices across numerous prominent vendors. In Figure 4, the graph displays the cumulative count of devices sourced from the database. This visualization offers a clear depiction of the overall volume of devices included in the analysis. It acts as a fundamental point of reference for comprehending the volume of data being examined in the research.

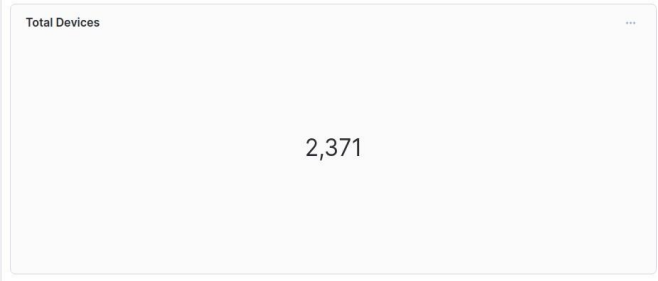

Figure 6: The Snowflake Scoreboard for Total Devices Count.

The analysis reveals a substantial number of devices that remain in use well past their EoL dates, with over 1000 devices reached EOL for over five years. The Snowflake dashboard proved essential in visualizing and managing this data, providing actionable insights that enable organizations to prioritize security measures for vulnerable devices and refine IT asset management practices, thereby enhancing both compliance and security frameworks within organizations.

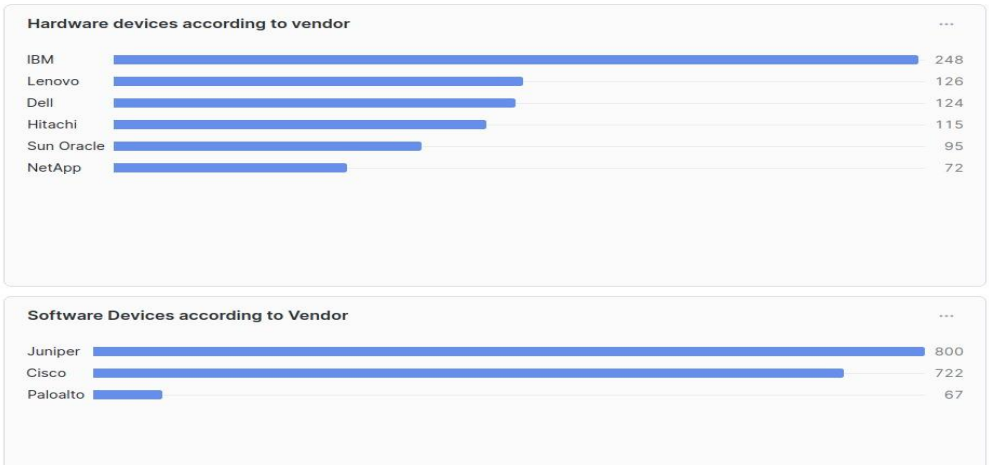

Figure 7: The Snowflake Graph for Devices according to Vendor.

Chapter 5

CONCLUSION

5. CONCLUSION

In conclusion, the "Effective End-of-Life (EOL) Management in Organization" project represents a transformative milestone for our organization, fostering a proactive approach to address vulnerabilities and adeptly manage the dynamic landscape of changing EOL statuses. The implementation of advanced strategies, including the scraping mechanism for data collection and the development of a centralized, scalable, and reliable dashboard, has yielded tangible results in providing stakeholders with unified, data-driven insights. The positive impact of this project extends across various dimensions of our organizational framework. Operational efficiency has markedly improved, enabling us to navigate the complexities of device management with greater ease. The enhanced data security measures implemented as part of this initiative contribute to a robust cybersecurity posture, instilling confidence among stakeholders and ensuring compliance with industry standards. Furthermore, the newfound agility in decision-making positions our organization to swiftly adapt to evolving industry demands. The scalability inherent in our approach ensures that our systems and processes can accommodate growth and changes seamlessly. As a result, our organization is better positioned to meet the challenges of the dynamic industrial landscape, with a future-oriented outlook that is both secure and efficient.

In essence, the "Effective End-of-Life (EOL) Management in Organization" project stands as a testament to our commitment to innovation and excellence in device management. It marks a significant stride toward a more secure, efficient, and informed future, setting the stage for continued success and resilience in the face of emerging challenges within the realm of device security and operational stability.

Chapter 6

FUTURE SCOPE

6. FUTURE SCOPE

The future scope for the project "Effective End-of-Life Management of Assets in Organizations by Snowflake Integration" is vast and encompasses several potential avenues for further development and research.

The project can expand to incorporate for efficient EOL management and real-time monitoring to extend asset life. Additionally, adapting the system for compliance with evolving international standards will help organizations stay compliant and efficient. Exploring these areas will significantly improve asset management practices and operational efficiencies across various industries. As digital transformation in EOL management becomes more widespread, there may be a need for regulatory frameworks and industry standards to ensure data privacy, security, and interoperability. Future research could contribute to the development of guidelines and best practices for digital EOL management, aligning with regulatory requirements and industry standards

REFERENCES

REFERENCES

[1]. Muyan Jiang, Ying Chen, Xin Chen, Javad Lavaei, and Anil Aswani "Optimal Contract Design for End-of-Life Care Payments" https://doi.org/10.48550/arXiv.2403.15099 arXiv:2403.15099 (math) 22 Mar 2024

[2]. Wayes Tushar, Dustin Niyato, Tapan K. Saha, H. Vincent Poor, and Chau Yuen "Blockchainenabled Circular Economy: Collaborative Responsibility in Solar Panel Recycling" https://arxiv.org/abs/2403.09937 15 Mar 2024

[3]. Craig Jacobik "Asset Ownership Identification: Using machine learning to predict enterprise asset ownership" 15 Dec 2023 https://arxiv.org/abs/2312.10266

[4]. Rohan Mohapatra "Large-scale End-of-Life Prediction of Hard Disks in Distributed Datacenters" 20 Mar 2023 in 2023 IEEE International Conference on Smart Computing (SMARTCOMP) https://doi.org/10.1109/SMARTCOMP58114.20 23.00069

[5]. Damjan MALETIČ1, Marta GRABOWSKA2, Matjaž MALETIČ1 "Drivers and Barriers of Digital Transformation in Asset Management" March 2023 Management and Production Engineering Review - pp. 118–126 DOI: 10.24425/mper.2023.145370

[6]. Extracting High-Quality Data Through Web Scraping: https://www.forbes.com/sites/forbestechcouncil/2 023/03/24/extracting-high-quality-data-through-web-scraping/?sh=583dfe4e8c67

[7]. Prakash, Vijay, Sicheng Xie, and Danny Yuxing Huang. "Software Update Practices on Smart Home IoT Devices." SCORED'22: Proceedings of the 2022 ACM Workshop on Software Supply Chain Offensive Research and

[8]. Ecosystem Defenses November 2022Pages 93–103 https://doi.org/10.1145/3560835.3564551

[9]. Mike McNeil Forbes Councils Member “How to Prepare Your Organization for The Future of Device Management” Forbes Technology Council Aug 29, 2022

[10]. Shweta More and Priya Daniel “AN ANALYSIS OF SNOWFLAKE: A PIVOTAL TOOL IN THE FIELD OF DATA ANALYSIS” IJFANS International Journal of Food and Nutritional Sciences Volume 11, Issue 10 (2022)

[11]. IEEE: A. S. Bale, N. Ghorpade, R. S, S. Kamalesh, R. R and R. B. S, "Web Scraping Approaches and their Performance on Modern Websites" 2022 3rd International Conference on Electronics and Sustainable Communication Systems (ICESC), Coimbatore, India, 2022, pp. 956-959, doi: 10.1109/ICESC54411.2022.9885689.

[12]. Dr Zena Assaad and Dr Mina Henein “End-ofLife of Software How is it Defined and Managed?” White paper April 2022

[13]. Emin Ozyoruk, Nesim Kohen Erkip, C¸ a˘gın Ararat “End-of-Life Inventory Management Problem: Results and Insights” arXiv:2101.09729v2 [math. OC] 26 Oct 2021

[14]. Ikechukwu Onyenwe “Developing Products Update-Alert System for e-Commerce Websites Users Using HTML Data and Web Scraping Technique” International Journal on Natural Language Computing 2 Sep 2021 https://doi.org/10.48550/arXiv.2109.00656

[15]. Wang, Dingding, et al. “A Measurement Study on the (In) security of End-of-Life (EoL) Embedded Devices.” arXiv preprint arXiv:2105.14298 (2021).

[16]. Sunny Mehta, Prof. Gayatri Pandi (Jain) "An Improving Approach for Fast Web Scrapping Using Machine Learning and Selenium Automation" International Journal of Advanced Research in Computer Engineering & Technology (IJARCET) Volume 8, Issue 10, October 2019, ISSN: 2278 – 1323

[17]. The Snowflake Documentation: https://docs.snowflake.com

DISSEMINATION OF WORK

PUBLICATION DETAILS

Paper Title	Journal Name	Issue Date	ISSN No. & DOI
Effective End-of-Life Management of Assets in Organization by Snowflake Integration	**IJARSCT** International Journal of Advanced Research in Science, Communication and Technology	4 April 2024	2581-9429 10.48475/IJARSCT-17450


INTERNATIONAL JOURNAL OF ADVANCED RESEARCH IN SCIENCE, COMMUNICATION AND TECHNOLOGY
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
IJARSCT
CERTIFICATE OF PUBLICATION
INTERNATIONAL STANDARD SERIAL NUMBER ISSN NO: 2581-9429
THIS IS TO CERTIFY THAT
Surabhi Lahoti
Shri Sant Gajanan Maharaj College of Engineering, Shegaon, Maharashtra, India
HAS PUBLISHED A RESEARCH PAPER ENTITLED
Effective End-of-Life Management of Assets in Organization by Snowflake Integration
IN IJARSCT, VOLUME 4, ISSUE 4, APRIL 2024
Certificate No: 042024-A1257
www.ijarsct.co.in
Crossref
DOI: 10.48175/IJARSCT-17450
www.doi.org
www.crossref.org
7.53
www.rpri.com
Editor-in-Chief


INTERNATIONAL JOURNAL OF ADVANCED RESEARCH IN SCIENCE, COMMUNICATION AND TECHNOLOGY
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
IJARSCT
CERTIFICATE OF PUBLICATION
INTERNATIONAL STANDARD SERIAL NUMBER ISSN NO: 2581-9429
THIS IS TO CERTIFY THAT
Kunal Chandore
Shri Sant Gajanan Maharaj College of Engineering, Shegaon, Maharashtra, India
HAS PUBLISHED A RESEARCH PAPER ENTITLED
Effective End-of-Life Management of Assets in Organization by Snowflake Integration
IN IJARSCT, VOLUME 4, ISSUE 4, APRIL 2024
Certificate No: 042024-A1259
www.ijarsct.co.in
Crossref
DOI: 10.48175/IJARSCT-17450
www.doi.org
www.crossref.org
7.53
www.rpri.com
Editor-in-Chief


INTERNATIONAL JOURNAL OF ADVANCED RESEARCH IN SCIENCE, COMMUNICATION AND TECHNOLOGY
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
IJARSCT
CERTIFICATE OF PUBLICATION
INTERNATIONAL STANDARD SERIAL NUMBER
ISSN NO: 2581-9429
THIS IS TO CERTIFY THAT
Aditya More
Shri Sant Gajanan Maharaj College of Engineering, Shegaon, Maharashtra, India
HAS PUBLISHED A RESEARCH PAPER ENTITLED
Effective End-of-Life Management of Assets in Organization by Snowflake Integration
IN IJARSCT, VOLUME 4, ISSUE 4, APRIL 2024
Certificate No: 042024-A1258
www.ijarsct.co.in
Crossref
DOI: 10.48175/IJARSCT-17450
www.doi.org
www.crossref.org
7.53
www.rpri.com
Editor-in-Chief

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 4, April 2024

Effective End-of-Life Management of Assets in Organization by Snowflake Integration

Dr. J. M. Patil[1], Prof. N. N. Ghuikar[2], Surabhi Lahoti[3], Aditya More[4], Kunal Chandore[5]
Assoc. Professor, Department of Computer Science and Engineering[1]
Asst. Professor, Department of Computer Science and Engineering[2]
Students, Department of Computer Science and Engineering[3,4,5]
Shri Sant Gajanan Maharaj College of Engineering, Shegaon, Maharashtra, India
jaimpatil1011@gmail.com, ghuikarnayana@gmail.com, sglohati123@gmail.com
yashmore1209@gmail.com, kunalachandore@gmail.com

Abstract: *To preserve device stability and security in today's industrial environment, vulnerabilities must be found and fixed. Asset managers face issues in managing End-of-Life (EOL) statuses, which are subject to constant change, and in traversing many data sources. It is challenging to have uniform view into important data without a centralized dashboard.*

We suggest putting in place a scraping method to collect data from various sources in order to overcome these issues. After the EOL data has been thoroughly verified and updated, the results will be shown on a consolidated, scalable, and dependable dashboard that easily links into the current systems. Advanced data slicing and dicing techniques will be used, utilizing Snowflake's capabilities, to handle enormous datasets efficiently. This all-encompassing strategy seeks to improve asset management procedures, strengthen device security, and offer useful information for defensible decision-making

Keywords: End-of-Life (EOL), Centralized Dashboard, Data Slicing and Dicing, Asset Management, Snowflake

I. INTRODUCTION

In the modern interconnected landscape, the seamless management of devices has become an integral aspect of our everyday lives. It emphasizes the importance of effectively managing device vulnerabilities, navigating through evolving End-of-Life (EOL) statuses, and consolidating data from various sources. The absence of a centralized dashboard is identified as a critical factor contributing to significant operational challenges in addressing these issues. This establishes the context for the proposed solution, suggesting that the implementation of a centralized dashboard and a scraping mechanism will be instrumental in overcoming these hindrances and improving operational efficiency.

In the ever-evolving landscape of the contemporary industry, the effective management of device vulnerabilities, dynamic adjustments to shifting End-of-Life (EOL) statuses, and the consolidation of data from a myriad of sources pose formidable challenges. These challenges are exacerbated by the notable absence of a centralized dashboard, thereby impeding operational efficiency and decision-making processes. In response to this complex scenario, the overarching goal of this project is to devise a comprehensive solution. This solution will not only navigate the intricacies of managing device vulnerabilities and adapting to EOL fluctuations but also tackle the data consolidation dilemma through the implementation of cutting-edge web scraping techniques.

At the heart of this endeavor lies the vision of creating a centralized, scalable, and reliable dashboard, meticulously designed to seamlessly integrate into existing infrastructures. This dashboard is poised to revolutionize the data visualization landscape, offering a holistic perspective on critical information. To fortify the infrastructure handling vast and diverse datasets, our

approach incorporates advanced data slicing and dicing techniques, harnessing the power of Snowflake, a leading data warehousing platform. Through the amalgamation of these innovative strategies, this project aspires to not only surmount existing challenges but also set a precedent for elevating operational efficiency, device security, and decision-making paradigms within the industry.

DOI: 10.48175/IJARSCT-17450

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 4, April 2024

1.1. End-Of-Life (EOL)

End-of-life (EOL) is a vital component in the expected lifespan of technologies and products. It denotes the stage at which a product or technology reaches the end of its commercially viable existence, rendering it obsolete or discontinued. In the context of devices and technologies, EOL signifies that the manufacturer ceases to produce, support, or update the product, thereby impacting its usability, security, and overall functionality. Managing the transition and implications of EOL is crucial in industries reliant on technological infrastructure to ensure continued operational efficiency and security. This introduction sets the stage for understanding the challenges associated with changing EOL statuses and the importance of addressing them in the proposed comprehensive solution.

1.2. Data sources

The landscape of data sources in the contemporary industry is characterized by its diversity and complexity. In the realm of device security and vulnerability management, the challenge lies in effectively harnessing information from a multitude of sources. This includes staying abreast of changing End-of-Life (EOL) statuses, understanding device vulnerabilities, and consolidating data for a comprehensive overview. However, this task is further complicated by the absence of a centralized dashboard, which results in a fragmented understanding of critical data.

In response to this, our proposed solution involves the strategic implementation of a scraping mechanism, designed to efficiently gather pertinent data from diverse sources. This initiative aims to streamline the process of data verification and EOL status updates, ultimately leading to the creation of a centralized, scalable, and reliable dashboard. By addressing the intricacies of data consolidation and management, this approach seeks to enhance overall visibility and understanding of crucial information, thus fortifying the foundation for improved device security and operational stability.

1.3 Asset management

Within the contemporary industrial landscape, asset management stands as a linchpin for organizational success. In the realm of device security and operational stability, the effective management of assets is paramount. The challenges arise from the intricate task of identifying and mitigating vulnerabilities, navigating through dynamic End-of-Life (EOL) statuses, and consolidating data from diverse sources. Compounded by the absence of a centralized dashboard, these challenges hinder the seamless management of assets, adding layers of complexity to the entire process.

To address these issues, a proposed solution involves the implementation of a scraping mechanism designed to collect data from multiple sources. Following rigorous verification and EOL updates, the outcomes are presented through a centralized, scalable, and reliable dashboard. This dashboard, designed for seamless integration, becomes the cornerstone for streamlined asset management. Leveraging advanced data slicing and dicing techniques using Snowflake, this approach aims to revolutionize asset management practices, providing a comprehensive solution to the multifaceted challenges faced in today's dynamic industrial landscape.

II. LITERATURE SURVEY

Ensuring compliance and receiving real-time updates on device status are pivotal aspects of maintaining organizational security, especially in light of recent findings from our company's report. The report indicates that a significant portion of security teams are tasked with overseeing an extensive number of devices, with a quarter managing upwards of 250,000 devices and over 10,000 workstations. However, despite the critical nature of device management, only 23% of security teams have fully enrolled their organization's devices in their Master Data Management (MDM) solution. Additionally, merely 25% have ensured that all or nearly all of their devices are running the latest operating system. These statistics underscore the challenges posed by current MDM solutions, which often lack comprehensive functionality, impeding security teams ability to maintain full visibility over devices, ensure compliance, and monitor devices in real-time. In particular, existing MDM solutions frequently provide incomplete data, leading to an inadequate understanding of an organization's device landscape. Furthermore, the absence of feedback mechanisms means that changes made by MDM solutions often go unnoticed, further complicating the task of device management. [8]

In the era of digital transformation, often dubbed the information age, there has been a pervasive shift towards digital technologies. This transition has led to the proliferation of intangible artifacts and services that are increasingly

DOI: 10.48175/IJARSCT-17450

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 4, April 2024

integrated into our daily lives. Unlike physical artifacts, intangible systems are immune to the typical wear and tear associated with material degradation, such as corrosion. However, despite this advantage, software still possesses a finite lifespan. Identifying the point at which a software product reaches its End-of-Life (EOL) is a multifaceted challenge. Currently, there is a scarcity of literature addressing practices for determining the EOL of software. Moreover, there exists a notable gap in the literature concerning explored or proposed methodologies for decommissioning software. [11]

The internet serves as a vast repository of information, catering to diverse needs ranging from business to personal use. Web Scraping (WS) emerges as a pivotal method for extracting substantial amounts of data from websites, typically in the form of unstructured HTML, and transforming it into structured data for various applications. While web scraping itself is not inherently illegal, ethical considerations often come into play. This paper delves into research findings regarding different web scraping techniques employed to extract data from websites, utilizing approaches such as the requests library, selenium, and other external libraries. The categorization of results sheds light on the susceptibility of different website categories to bot attacks and highlights variations in security levels across categories. To enhance result accuracy, the paper suggests identifying additional modules or libraries for similar tests and conducting tests on a larger scale. Additionally, the effectiveness of selenium-based tests can be further assessed by incorporating additional parameters or employing multiple parameter checks. [10]

The literature surrounding the utilization of Selenium Automation for purposes beyond testing predominantly focuses on its efficacy in web scraping and data extraction tasks. While Selenium has traditionally been employed for error detection and testing during software development, recent studies have explored its potential for automating data collection processes from web pages. Techniques such as Count Vectorizer Machine Learning and Forward Selection Wrapper methods of Feature Selection have been integrated with Selenium to enhance its capabilities in identifying and extracting relevant information from web documents. However, existing research also highlights challenges such as potential inaccuracies in data extraction and the need for manual intervention to address errors. Nonetheless, the demonstrated versatility of Selenium Automation beyond testing underscores its value in streamlining data retrieval tasks, paving the way for applications in various domains such as marketing and client services. [15]

In the contemporary business landscape, data reigns supreme, serving as the cornerstone for strategic decision-making across all industries and domains. Recognized as one of the most valuable assets, data empowers stakeholders to make informed choices through meticulous analysis. However, the sheer volume of data generated necessitates tools with attributes such as high performance, scalability, and user-friendliness, enabling businesses to extract maximum value in a cost and time-efficient manner.

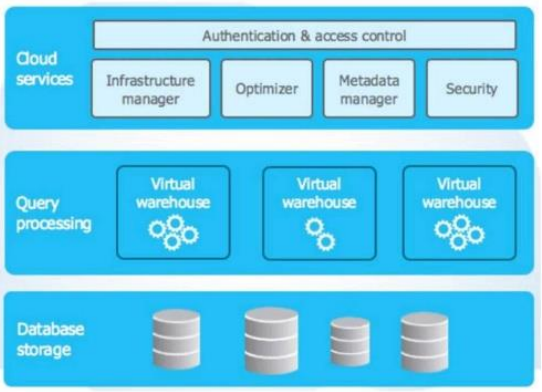

Figure 1: Snowflake Architecture

Amidst the era of distributed computing, a plethora of cloud platforms and Software as a Service (SaaS) providers offer unlimited computation and storage resources on demand. This paper delves into Snowflake Elastic Data Warehouse, a prominent tool utilized by organizations to tackle real-world data challenges, such as managing vast datasets with

DOI: 10.48175/IJARSCT-17450

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 4, April 2024

exceptional performance. Examining Snowflake's architecture, data sharing and storage capabilities, as well as its data ingestion, transformation, analytics, and visualization features, this paper elucidates why Snowflake stands out as the preferred solution for handling substantial volumes of data with rapid performance. [9]

End-of-Life (EoL) devices marking a significant contribution to the existing literature. Through a meticulous analysis comprising aliveness and vulnerability assessments, the study unveils critical insights previously unknown to the research community. By examining a substantial number of EoL models from prominent vendors such as D-Link, Tp-Link, and Netgear over a ten-month period, the study sheds light on the prevalence and persistence of active EoL devices. Notably, the findings reveal a startling revelation of over 2 million active EoL devices, with nearly 300,000 still operational even five years post-EoL status. Moreover, the study highlights a concerning trend where a significant portion of vulnerabilities, totalling 182 out of 294, are identified after the EoL date, indicating a lapse in post-EoL security patching practices by vendors. Consequently, the study underscores the urgent need for a more structured and comprehensive approach to addressing the security risks posed by active EoL devices, with over 1 million devices vulnerable to high-risk exploits, posing a significant threat to cybersecurity landscapes.[14]

The research conducted by Wang et al. and Jiang et al. represents a pioneering effort in investigating the (in)security of End-of-Life embedded devices. The aliveness analysis and vulnerability analysis provide a holistic view of the challenges posed by active EoL devices, contributing to the body of knowledge in the field of cyber security. This literature review highlights the importance of their work, emphasizing the need for continued research and collaborative efforts to address the evolving threats associated with EoL devices.[14]

Asset owner identification is paramount for information security organizations, aiding in breach detection, vulnerability assessment, and countermeasure definition. This research employs various machine learning algorithms to predict asset ownership, conducting separate analyses for each owner and utilizing a 100-iteration Monte Carlo Cross Validation. The resulting visualization dashboard enables exploratory data analysis and model evaluation. Adaboost emerges as the top-performing model, while Naïve Bayes lags. Key features like Fully Qualified Domain Name (FQDN), CIDR, and location significantly impact ownership prediction.[3]

The proliferation of renewable energy sources, notably solar power, has surged in recent years. However, the consequential environmental risks stemming from excessive solar panel installations and inadequate recycling infrastructure present pressing challenges. This paper advocates for a circular economy approach to mitigate these issues, proposing the integration of blockchain technology to track the end-of-life (EOL) of solar panels and assign responsibilities to stakeholders. Furthermore, the paper suggests monetizing panel degradation by monitoring users' energy-related activities and utilizing the generated funds for future recycling efforts. Introducing a novel cryptocurrency, the recycling coin (RC-coin), serves as an incentive for solar panel recycling, while leveraging decentralized finance mechanisms to address coin price stability and supply concerns. [2]

The paper aims to enhance comprehension of the catalysts and impediments to digital transformation within asset management. Employing a qualitative Delphi study involving 15 experts from academia, consultancy, and industry sectors, the research endeavors to discern, validate, and categorize the drivers and barriers influencing digital transformation in asset management. Through expert interactions, the study identifies 20 barriers, encompassing factors such as strategic misunderstandings, lack of vision or strategy, cultural inertia, deficient asset management systems, limited awareness of digital trends, and inadequacies in employee knowledge and skills. Additionally, the study elucidates 12 critical drivers essential for the digital transformation of asset management, including cost reduction, opportunities in condition monitoring, and anticipated benefits in risk management processes. The findings from this research offer valuable insights for organizations contemplating digital transformation initiatives in asset management, emphasizing the necessity of addressing both barriers and drivers to effectively navigate the transformative journey.[5]

The evolution of technology has revolutionized traditional business practices, particularly with the emergence of e-commerce websites as prominent platforms for conducting transactions. This transition has significantly impacted the dynamics among marketers, retailers, and consumers, collectively referred to as users in this context, by providing convenient avenues for buying and selling goods online. The proliferation of e-commerce websites has simplified product search operations, facilitating tasks such as accessing updated information on new or edited products, sales, and costs. Without employing methods like the one presented in this paper or those discussed in existing literature, navigating through e-commerce marketplaces for activities like price comparison would be time-consuming. Given the

DOI: 10.48175/IJARSCT-17450

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 4, April 2024

rapidly changing nature of prices and product information, it is imperative to analyze competitor prices and intermediary sellers' roles regularly to make informed decisions. [13]

In our project, we're leveraging cutting-edge technology to address the complex challenges associated with managing End-of-Life (EOL) assets in IT Asset Management (ITAM). Specifically, we're focusing on utilizing web scraping techniques to automate the extraction of EOL dates directly from manufacturers' websites and digital platforms. By doing so, we aim to streamline the acquisition of critical EOL data, significantly improving its accuracy and timeliness. However, implementing web scraping at scale introduces certain complexities related to data integrity, scalability, and legal compliance. To tackle these challenges, we're developing a robust framework that ensures the reliability and regulatory compliance of the collected data.

Furthermore, our project involves integrating Snowflake's cloud-based data warehousing solution into our workflow. Snowflake is renowned for its scalability, performance, and user-friendly interface, making it an ideal partner for handling the vast amounts of EOL data collected through web scraping. By leveraging Snowflake's capabilities in data storage, processing, and analytics, we aim to create a centralized repository for EOL information. This integration will not only facilitate efficient storage of data but also enable us to derive actionable insights for forecasting asset renewals and prioritizing security measures for assets nearing or surpassing their EOL.

Ultimately, our project aims to go beyond enhancing operational efficiency. We see effective EOL asset management as a critical component of bolstering an organization's digital security posture. By reallocating resources from maintaining outdated assets to adopting modern technologies, we anticipate significant improvements in mitigating the risks of cyber threats and regulatory non-compliance. Our goal is to optimize IT budgets while enhancing overall security measures, ultimately enhancing enterprises' resilience and effectiveness in the current digital environment.

III. METHODOLOGY

The first step involves identifying a comprehensive list of End-of-Life (EoL) models from various vendors. This selection process includes prominent vendors such as Lenovo, Dell, Hitachi among others. EoL models are chosen based on their prevalence and significance in the market, ensuring a representative sample for analysis. Web scraping is employed to extract relevant data regarding EoL dates and device information from the official websites of selected vendors. Selenium, a powerful automation tool, is utilized for web scraping due to its capability to interact with dynamic web elements. Selenium scripts are developed to navigate through vendor websites, locate pertinent data, and extract it systematically.

3.1. Web Scraping Mechanism:

Web scraping is the process of extracting data from websites. The methodology involves:

- Data Collection: Identifying the target websites and defining the data to be extracted. This includes understanding the website structure and elements.
- Use of Libraries: Utilizing tools like Selenium, Beautiful Soup, or Scrapy in Python for efficient data extraction.
- Selenium is often used for dynamic content loading through automation.
- Programming Languages: Python is a commonly used language due to its rich ecosystem of web scraping libraries and tools. It allows for handling HTTP requests, parsing HTML, and automating browser interactions.

3.2. Data Storage and Management:

After scraping, the data needs to be stored and managed effectively. The methodology includes:

- Database or Data Warehouse: Choosing an appropriate database or data warehouse for storage. Options include SQL databases (like MySQL or PostgreSQL), NoSQL databases (like MongoDB), or cloud-based data warehouses (like Amazon Redshift, Google Big Query).
- Data Cleansing: Ensuring data quality by identifying and correcting errors or inconsistencies. This involves handling missing values, removing duplicates, and standardizing formats.

DOI: 10.48175/IJARSCT-17450

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 4, April 2024

3.3. Snowflake Integration:

Integrating with Snowflake, a cloud-based data platform, involves several steps:

- Data Loading: Loading the scraped and cleansed data into Snowflake. This may include using Snowflake connectors or ETL (Extract, Transform, Load) tools.
- Data Processing: Performing necessary transformations on the data within Snowflake to prepare it for analysis. This could involve aggregations, joins, or other operations.
- Data Transformation: Adapting the data to meet specific reporting or analysis requirements.

3.4. Scalability and Performance Optimization:

Ensuring the system can handle growth and optimizing performance involves:

- Database Scaling: Scaling the database infrastructure to handle increased data volumes. This may involve vertical scaling (upgrading hardware) or horizontal scaling (adding more servers).
- Query Optimization: Optimizing database queries to reduce response times. This includes indexing, query caching, and other performance-tuning techniques.
- Dashboard Performance: Optimizing the performance of the visualization tools and dashboards. This may involve caching, aggregations, and efficient use of resources.

3.5. Data Security and Privacy:

Ensuring the security and privacy of the data involves:

- Access Control: Implementing role-based access control (RBAC) to restrict access to sensitive data based on user roles.
- Compliance: Adhering to data protection regulations and industry-specific compliance standards. This may involve encryption, anonymization, and audit trails.
- 3.6. Centralized Dashboard:
- Creating a centralized dashboard involves:
- User Interface (UI): Designing an intuitive and user-friendly interface for the dashboard.
- Visualization: Choosing appropriate data visualization techniques to effectively communicate insights.
- Filters and Interactivity: Implementing filters and interactive elements to allow users to explore and analyze the data dynamically.

IV. RESULTS

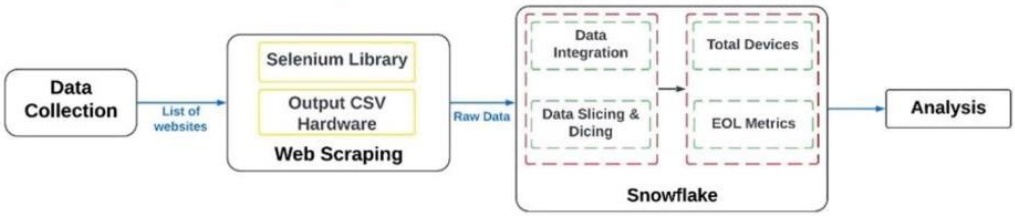

Figure 2: The approach of our study.

This section presents visual representations of the results obtained from the tests conducted. These graphs serve to enhance the comprehension of the article's objectives by providing clear and accessible insights into the data analysis performed.

The seamless integration of Selenium web scraping and Snowflake data management effectively pinpointed more than 2,000 active End-of-Life (EoL) devices across numerous prominent vendors. In Figure 2, the graph displays the cumulative count of devices sourced from the database. This visualization offers a clear depiction of the overall volume of devices included in the analysis. It serves as a foundational reference point for understanding the scale of data under examination in the study.

DOI: 10.48175/IJARSCT-17450

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 4, April 2024

Figure 3: Total Devices in database

The analysis reveals a substantial number of devices that remain in use well past their EoL dates, with over 1000 devices reached EOL for over five years. The Snowflake dashboard proved essential in visualizing and managing this data, providing actionable insights that enable organizations to prioritize security measures for vulnerable devices and refine IT asset management practices, thereby enhancing both compliance and security frameworks within organizations

V. CONCLUSION

In conclusion, the "Effective End-of-Life (EOL) Management in Organization" project represents a transformative milestone for our organization, fostering a proactive approach to address vulnerabilities and adeptly manage the dynamic landscape of changing EOL statuses. The implementation of advanced strategies, including the scraping mechanism for data collection and the development of a centralized, scalable, and reliable dashboard, has yielded tangible results in providing stakeholders with unified, data-driven insights.

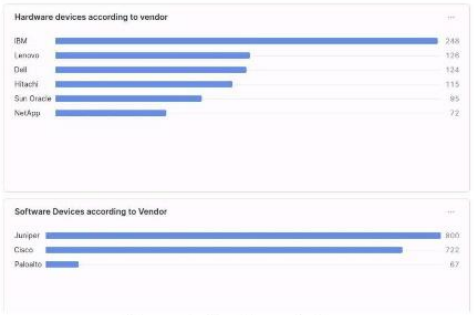

Figure 4: Dashboard view

The positive impact of this project extends across various dimensions of our organizational framework. Operational efficiency has markedly improved, enabling us to navigate the complexities of device management with greater ease. The enhanced data security measures implemented as part of this initiative contribute to a robust cybersecurity posture, instilling confidence among stakeholders and ensuring compliance with industry standards.

Furthermore, the newfound agility in decision-making positions our organization to swiftly adapt to evolving industry demands. The scalability inherent in our approach ensures that our systems and processes can accommodate growth and changes seamlessly. As a result, our organization is better positioned to meet the challenges of the dynamic industrial landscape, with a future-oriented outlook that is both secure and efficient.

In essence, the "Effective End-of-Life (EOL) Management in Organization" project stands as a testament to our commitment to innovation and excellence in device management. It marks a significant stride toward a more secure, efficient, and informed future, setting the stage for continued success and resilience in the face of emerging challenges within the realm of device security and operational stability.

REFERENCES

[1] Muyan Jiang, Ying Chen, Xin Chen, Javad Lavaei, and Anil Aswani "Optimal Contract Design for End-of-Life Care Payments" https://doi.org/10.48550/arXiv.2403.15099 arXiv:2403.15099 (math) 22 Mar 2024

DOI: 10.48175/IJARSCT-17450

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 4, April 2024

[2]Wayes Tushar, Dustin Niyato, Tapan K. Saha, H. Vincent Poor, and Chau Yuen "Blockchain-enabled Circular Economy: Collaborative Responsibility in Solar Panel Recycling" https://arxiv.org/abs/2403.09937 15 Mar 2024

[3]Craig Jacobik "Asset Ownership Identification: Using machine learning to predict enterprise asset ownership" 15 Dec 2023 https://arxiv.org/abs/2312.10266

[4]Rohan Mohapatra "Large-scale End-of-Life Prediction of Hard Disks in Distributed Datacenters" 20 Mar 2023 in 2023 IEEE International Conference on Smart Computing (SMARTCOMP) https://doi.org/10.1109/SMARTCOMP58114.2023.00069

[5]Damjan MALETIČ1, Marta GRABOWSKA2, Matjaž MALETIČ1 "Drivers and Barriers of Digital Transformation in Asset Management" March 2023 Management and Production Engineering Review - pp. 118–126 DOI: 10.24425/mper.2023.145370

[6]Extracting High-Quality Data Through Web Scraping: https://www.forbes.com/sites/forbestechcouncil/2023/03/24/extracting-high-quality-data-through-web-scraping/?sh=583dfe4e8c67

[7]Prakash, Vijay, Sicheng Xie, and Danny Yuxing Huang. "Software Update Practices on Smart Home IoT Devices." SCORED'22: Proceedings of the 2022 ACM Workshop on Software Supply Chain Offensive Research and Ecosystem Defenses November 2022Pages 93–103 https://doi.org/10.1145/3560835.3564551

[8]Mike McNeil Forbes Councils Member "How to Prepare Your Organization for The Future of Device Management" Forbes Technology Council Aug 29, 2022

[9]Shweta More and Priya Daniel "AN ANALYSIS OF SNOWFLAKE: A PIVOTAL TOOL IN THE FIELD OF DATA ANALYSIS" IJFANS International Journal of Food and Nutritional Sciences Volume 11, Issue 10 (2022)

[10]IEEE: A. S. Bale, N. Ghorpade, R. S, S. Kamalesh, R. R and R. B. S, "Web Scraping Approaches and their Performance on Modern Websites" 2022 3rd International Conference on Electronics and Sustainable Communication Systems (ICESC), Coimbatore, India, 2022, pp. 956-959, doi: 10.1109/ICESC54411.2022.9885689.

[11]Dr Zena Assaad and Dr Mina Henein "End-of-Life of Software How is it Defined and Managed?" White paper April 2022

[12]Emin Ozyoruk, Nesim Kohen Erkip, C¸ a˘gın Ararat "End-of-Life Inventory Management Problem: Results and Insights" arXiv:2101.09729v2 [math.OC] 26 Oct 2021

[13]Ikechukwu Onyenwe "Developing Products Update-Alert System for e-Commerce Websites Users Using HTML Data and Web Scraping Technique" International Journal on Natural Language Computing 2 Sep 2021 https://doi.org/10.48550/arXiv.2109.00656

[14]Wang, Dingding, et al. "A Measurement Study on the (In) security of End-of-Life (EoL) Embedded Devices." arXiv preprint arXiv:2105.14298 (2021).

[15]Sunny Mehta, Prof. Gayatri Pandi (Jain) "An Improving Approach for Fast Web Scrapping Using Machine Learning and Selenium Automation" International Journal of Advanced Research in Computer Engineering & Technology (IJARCET) Volume 8, Issue 10, October 2019, ISSN: 2278 – 1323

[16]The Snowflake Documentation: https://docs.snowflake.com

DOI: 10.48175/IJARSCT-17450

ISSN 2581-9429 IJARSCT

SPONSORSHIP DETAILS

Project Title	Category	Sponsored By
Effective End-of-Life Management in The Organization by Snowflake Integration	Industry Sponsored	ApexaiQ Technologies Pvt. Ltd.

APEXAIQ
Inventory | Security | Productivity | Finance

Date: 04th March 2024

TO WHOM SO IT MAY CONCERN

Subject: Sponsorship Letter

Dear Surabhi Lahoti,

We are pleased to inform you that ApexaiQ Technologies Pvt Ltd has chosen you to receive the sponsorship for your academic pursuits.

Below are the details:

- **Project Name:** Effective End-of-life Management in The Organization by Snowflake Integration
- **Duration:** This internship will cover the student beginning on [August'2023 to September'2023] and [January'2024 to March'2024].
- **Stipend:** You will receive a stipend of INR 7000 per month.

We look forward to witnessing your continued success and are excited to play a part in your educational journey.

Sincerely,

For, ApexaiQ Technologies Pvt Ltd.

Pragti Aggarwal

Pragti Aggarwal

Director

APEXAIQ TECHNOLOGIES PRIVATE LIMITED
Reg Add: F 78 Venus Appt. Sector 9 Rohini, Near Japanese Park, Delhi, India, 110085
CIN: U72900DL2022FTC406916
www.apexaiq.com

Doc ID: c72e89f1ccb815af79e141114172cdb12b70efbe

APEXAIQ
Inventory | Security | Productivity | Finance

Date: 04th March 2024

TO WHOM SO IT MAY CONCERN

Subject: Sponsorship Letter

Dear Kunal Chandore,

We are pleased to inform you that ApexaiQ Technologies Pvt Ltd has chosen you to receive the sponsorship for your academic pursuits.

Below are the details:

- **Project Name:** Effective End-of-life Management in The Organization by Snowflake Integration
- **Duration:** This internship will cover the student beginning on [August'2023 to September'2023] and [January'2024 to March'2024].
- **Stipend:** You will receive a stipend of INR 7000 per month.

We look forward to witnessing your continued success and are excited to play a part in your educational journey.

Sincerely,

For, ApexaiQ Technologies Pvt Ltd.

Pragti Aggarwal

Pragti Aggarwal

Director

APEXAIQ TECHNOLOGIES PRIVATE LIMITED
Reg Add: F 78 Venus Appt. Sector 9 Rohini, Near Japanese Park, Delhi, India, 110085
CIN: U72900DL2022FTC406916
www.apexaiq.com

Doc ID: 19a861e5a0527adb9319480099f1049ab01913cb

APEXAIQ
Inventory | Security | Productivity | Finance

Date: 04th March 2024

TO WHOM SO IT MAY CONCERN

Subject: Sponsorship Letter

Dear Aditya More,

We are pleased to inform you that ApexaiQ Technologies Pvt Ltd has chosen you to receive the sponsorship for your academic pursuits.

Below are the details:

- **Project Name:** Effective End-of-life Management in The Organization by Snowflake Integration
- **Duration:** This internship will cover the student beginning on [August'2023 to September'2023] and [January'2024 to March'2024].
- **Stipend:** You will receive a stipend of INR 7000 per month.

We look forward to witnessing your continued success and are excited to play a part in your educational journey.

Sincerely,

For, ApexaiQ Technologies Pvt Ltd.

Pragti Aggarwal

Pragti Aggarwal

Director

APEXAIQ TECHNOLOGIES PRIVATE LIMITED
Reg Add: F 78 Venus Appt. Sector 9 Rohini, Near Japanese Park, Delhi, India, 110085
CIN: U72900DL2022FTC406916
www.apexaiq.com

Doc ID: 7660c12a8a1ce23c74c9795e25db653790175fe1

Project Group Members

Name: Surabhi Lahoti

Branch: Computer Science and Engineering, SSGMCE

Address: Dhamangaon Rly, Maharashtra, India – (444709)

Email ID: sglahoti123@gmail.com

Mobile No.: +91 7262926190

Name: Kunal Chandore

Branch: Computer Science and Engineering, SSGMCE

Address: Buldhana, Maharashtra, India – (443001)

Email ID: kunalachandore@gmail.com

Mobile No.: +91 8329472891

Name: Aditya More

Branch: Information Technology, SSGMCE

Address: Shegaon, Maharashtra, India – (444203)

Email ID: yashmore1209@gmail.com

Mobile No.: +91 9421470122

Plagiarism Report (using Turnitin software)

EOL Research Paper.docx

ORIGINALITY REPORT

11%	10%	3%	2%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	ijircce.com Internet Source	4%
2	arxiv.org Internet Source	3%
3	Submitted to University of Greenwich Student Paper	1%
4	Submitted to Colorado Technical University Student Paper	<1%
5	Submitted to Westcliff University Student Paper	<1%
6	Xixing Li, Qiang Wei, Zehui Wu, Wei Guo, Linhao He. "Harden-IoT: hardening the EoL devices by intercepting the attack vector for future B5G/6G IoT", Wireless Networks, 2023 Publication	<1%
7	www.fivetran.com Internet Source	<1%
8	Dingding Wang, Muhui Jiang, Rui Chang, Yajin Zhou, Hexiang Wang, Baolei Hou, Lei Wu, Xiapu Luo. "An Empirical Study on the Insecurity of End-of-Life (EoL) IoT Devices", IEEE Transactions on Dependable and Secure Computing, 2024 Publication	<1%
9	ijritcc.org Internet Source	<1%
10	www.coursesidekick.com Internet Source	<1%
11	fastercapital.com Internet Source	<1%

Exclude quotes	Off	Exclude matches	Off
Exclude bibliography	On		