

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

Dr. J. M. Patil¹, Prof. N. N. Ghuikar², Surabhi Lahoti³, Aditya More⁴, Kunal Chandore⁵

Assoc. Professor, Department of Computer Science and Engineering¹

Asst. Professor, Department of Computer Science and Engineering²

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, sglohathi123@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.

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

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.

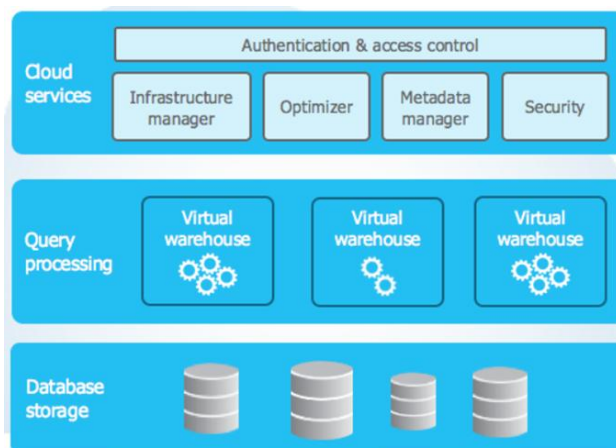


Figure 1: Snowflake Architecture

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

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

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, BeautifulSoup, 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.

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

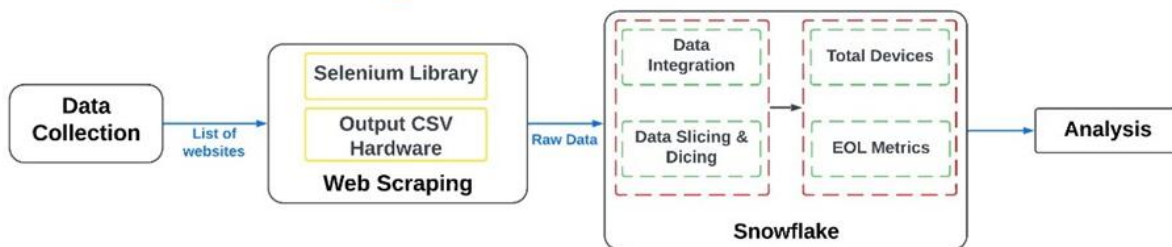


Figure 2: The approach of our study.

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.

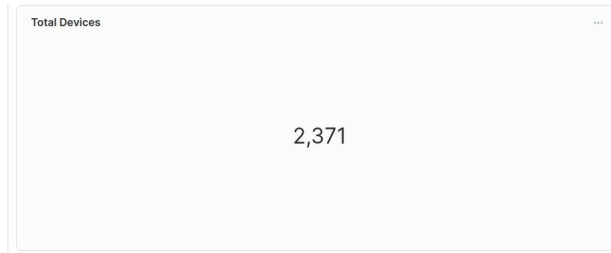


Figure 3: Total Devices in database

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.

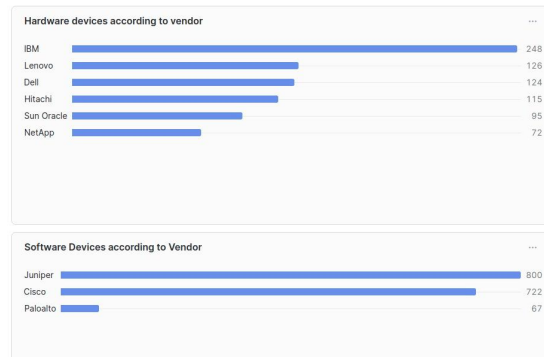


Figure 4: Dashboard view

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

- [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 ğın 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 Scraping 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>