PENERBIT UMT
Menjana Khazanah Ilmuwan

Journal of Mathematical Sciences and Informatics
Journal Homepage: https://journal.umt.edu.my/index.php/jmsi
eISSN: 2948-3697
DOI: http://doi.org/10.46754/jmsi.2023.12.006

JMSI
Journal of Mathematical Sciences and Informatics

THE IMPACT OF OPEN-SOURCE SOFTWARE ON ARTIFICIAL INTELLIGENCE

MYFTAHUDDIN HAZMI HASSRI* AND MUSTAFA MAN

Faculty of Ocean Engineering Technology and Informatics, Universiti Malaysia Terengganu, 21030 Kuala Nerus, Terengganu, Malaysia; hazmihassri@gmail.com, mustafaman@umt.edu.my

**Corresponding author: hazmihassri@gmail.com*

ARTICLE INFO

Article History:
Received
15 APRIL 2023
Accepted
1 OCTOBER 2023
Available online
25 JANUARY 2024

Section Editor:

Keywords:
Open-Source Software;
Proprietary Software;
Artificial Intelligence;
AI Libraries and Tools

ABSTRACT

Open-source software (OSS) is defined as applications or tools developed by a community. Richard Stallman, inspired by the liberty of using software for free, initiated the concept of OSS development. After successfully launching the GNU project, he established the Free Software Foundation to counter attempts by big technology companies to monopolise the tech economy. This effort led to the development and enduring usage of various OSS projects, including Apache and Linux OS. Presently, OSS spans numerous sectors, such as software development, 3D rendering, mobile applications, artificial intelligence (AI), and many more. This paper provides a review of OSS, encompassing its history, advantages and disadvantages, impact on industries and AI, and associated challenges.

INTRODUCTION

An Overview of Open-source Software

The concept of Open-source Software (OSS) revolves around the idea of open innovation in software development. OSS provides its source code and original rights, encouraging collaboration and knowledge sharing among programmers to enhance, modify, or fix aspects that need improvement. The fundamental principle behind OSS is that open software is more efficient than proprietary counterparts, as it promotes collaboration and knowledge exchange, resulting in the significant benefit of leveraging the knowledge and expertise of a larger open community for software production [1].

OSS offers free software developed by its community before being distributed to the public with certain conditions. Several traditional software companies derive benefits from free software by devising and utilising their business models within the OSS ecosystem. For instance, in the mobile industry, there have been multiple attempts to establish open-source societies for Android and Symbian projects. Utilising and relying on OSS are alternative approaches to minimising development costs while remaining competitive in the market [2].

In essence, OSS is free-licensed software that allows programmers to modify the software according to their needs and control its distribution. OSS has witnessed significant success and can compete with proprietary software in the market. For example, the Linux Operating System covers approximately 30% of the server operating system market, while Microsoft Operating System holds about 50% [3]. This paper aims to examine the impact of OSS on the information technology industry.

Background History

The roots of OSS trace back to the 1950s and 1960s when computer software was shared exclusively among developers and researchers with the goal of advancing technology. However, the concept of OSS as it is known today emerged in the late 1970s and early 1980s, a time when major companies monopolised the market with their proprietary software. Microsoft's Bill Gates, in his "Open Letter to Hobbyists", remarked that software is detrimental to the public, but not to private property [4]. Influential figures such as Gates, Ellison, Cullinane, and Goetz advocated for a robust software industry with closed code products [5].

These large companies sought to dominate the market to maximise profits. In response, software programmer Richard Stallman established the Free Software Foundation in 1985 after successfully launching the GNU project. This non-profit organisation introduced the concept of free software for users, emphasising their rights to use them. The foundation facilitated the distribution of free software along with its source code. The community greatly benefited from this initiative, as they could make modifications and redistribute their work without violating the copyright of the original software [6]. Stallman then initiated the GNU operating system, which comprised a set of tools that were free, modifiable, and redistributable to any user.

The GNU Project

The GNU project, founded by Stallman in 1983, is a collaborative effort in free software development. It's important to note that the term "free" in this context does not imply no cost, but rather underscores the freedom to use, understand, modify, share, and distribute the software [7]. The GNU Project focuses on four main aspects: users have the full capacity to utilize the software irrespective of their intentions or purposes; users can understand how the software works and modify it according to their needs; the software can be shared by anyone for the benefit of society; and, users can distribute modified versions for the community's benefit using the provided source code.

These principles are designed to enhance efficiency and effectiveness in software development, prevent fraud, ensure convenience, increase revenue, and reduce production costs. The GNU Project gained substantial popularity with the emergence of the Linux Operating System and Apache web server, swiftly dominating their respective markets. The development of OSS continued through the 2000s and beyond. The foundation had approximately 3,000 sponsor members and over 200,000 monthly newsletter subscribers in 2009 [8].

The surge in the OSS community members illustrates people's interest in understanding user rights and challenging proprietary software offered by large companies. Presently, there are more than 110,000 free software projects available, ranging from operating systems to software games [9]. Users have the option to choose alternative software instead of paying for proprietary counterparts, often with comparable features. Table 1 presents an example of some proprietary and OSS available in the market, along with their respective applications [10].

Table 1: Examples of OSS and proprietary software based on application

Application	Open Source	Proprietary
Web browser	Mozilla Firefox	Microsoft Internet Explorer, Netscape
Web server	Apache	Microsoft IIS, Netscape
Application server	Apache Tomcat, JBoss	WebMethods
Office suite	OpenOffice.Org	Microsoft Office, Corel WordPerfect Office
Email/collaboration	Ximian Evolution	Microsoft Exchange, Lotus Notes
Database	MySQL, PostgreSQL	IBM DB2, Oracle9i, MS SQL

WHAT IS ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING?

Artificial Intelligence (AI) refers to the capability of computers and machines to perform tasks that typically require human-like cognitive functions, such as speech recognition, decision-making, and visual perception. Realising AI involves the development of computer programmes and algorithms that can simulate human intelligence and learning processes. The applications of AI are vast, encompassing natural language processing, image and speech recognition, automation, and predictive analytics. As mentioned in [11], AI is rapidly advancing, and people are incorporating it into their daily lives.

Moreover, AI plays a crucial role in the development strategies of countries worldwide, enhancing competitiveness and ensuring national security. Its applications are diverse, ranging from precise task execution to data analysis for optimising operations such as agriculture, to improving decision-making processes and boosting production rates [12]. According to [13], AI holds immense potential for problem-solving, proving highly beneficial for both companies and non-profit organisations. It assists humans in making better decisions and provides insights into the future. The author highlights several benefits:

a) AI can complete highly complex tasks faster than humans with a higher success rate.

b) AI is less prone to errors and can perform numerous functions simultaneously.

c) AI can make predictions and explore new avenues of knowledge, contributing to the advancement of various fields.

The Relationship between OSS, AI and Machine Learning

OSS is freely accessible and can be modified without constraints. It plays a pivotal role in the development and utilisation of AI. OSS is employed in various programming languages, including Python, R, C++, and JavaScript, and is integral to implementing machine learning algorithms. Frameworks like Scikit-learn in Python have effectively streamlined a wide array of algorithms into a cohesive machine learning process, simplifying usage for data scientists in their daily tasks [14]. Moreover, OSS accelerates the adoption of AI. As highlighted in [15], the technical quality maintained by Open-Source Intelligence analysts is directly linked to AI system performance, showcasing OSS's potential to augment AI capabilities.

Advantages and Disadvantages of OSS

The advantages and disadvantages of OSS can be categorised into three main areas: users, developers, and systems [16].

Advantages

From a user's standpoint, OSS offers better flexibility. Users can import any library or tool from open archives and use them immediately, and gain insights into their functioning through community knowledge sharing. Additionally, OSS provides users with access to its source code, allowing for code modifications. For developers, OSS enables the creation of customisable solutions and the option to seek assistance from the community when faced with challenges. The reuse of existing code during tool or library modifications helps developers reduce their development time. In terms of system attributes, OSS is notably cost-effective and reliable, with a significant portion. OSS communities contribute to the rapid development of OSS systems, thanks to their vast number of development members.

Disadvantages

However, OSS systems are not without their shortcomings. Some users may encounter incomplete or bad documentation, making it a challenge to learn the installed tools or address errors during the installation process. The absence of proper documentation can pose difficulties for new developers the in understanding the OSS workflow. Moreover, OSS systems lack formal, process-centralised management, with no designated development team or single entity responsible for troubleshooting. This increases the likelihood of a particular OSS project being abandoned.

IMPACT OF OSS ON INDUSTRIES

Presently, OSS is widely used in many industries, revolutionising the landscape in numerous ways. OSS is typically free to use and distribute, making it an attractive choice for organisations looking to reduce licensing costs. This cost-saving advantage allows smaller businesses to compete more effectively with larger companies, narrowing the technology gap. Despite being termed as free software, certain OSS programmes may involve additional costs for support, services, or additional features, in contrast to the varying costs associated with closed-source software (CSS) depending on complexity [17].

The overarching concept is that open-source technologies will significantly impact various fields, such as Industry 4.0, machine learning, the Internet of Things (IoT), big data analytics, and cloud computing. In the current landscape, OSS plays an essential role in the development of IoT platforms and prototypes. Additionally, many machine learning engines are open source, enabling developers to experiment, collaborate, and learn from them [18].

AI Industry

OSS exerts a significant influence on the AI industry. Despite the dominance of proprietary software, the emergence of OSS has democratised access to powerful AI frameworks and libraries. This shift prevented the monopolisation of AI technology by large corporations, allowing individuals to experiment and enhance AI capabilities. Consequently, many OSS libraries, including TensorFlow, Keras, PyTorch, and Apache, have been developed and are extensively used in AI development, facilitating the creation of complex models and algorithms.

OSS enables and encourages AI integration by reducing the need for deep mathematical and technical understanding. The incorporation of intricate mathematical algorithms into code is known to be challenging and time-consuming. As noted in [19], OSS software like SciPy, Scilab, and GNU Octave, among others, serves as essential libraries l for numerical computing, aiding in computations of complex equations, plotting 2D and 3D functions, and performing data analytics.

The OSS community fosters collaboration and knowledge exchange, driving innovation in AI. For instance, the Apache Spark project, an open-source AI software and data processing engine, was collaboratively developed by the community based on existing code. This collaborative approach also contributes to reducing AI bias, as the decentralised nature of the codes—contributed by individuals from diverse organisations, backgrounds, and industries—makes it more likely for the OSS community to create new and innovative code compared with CSS communities found in large companies [20].

OSS is widely acknowledged for its potential to significantly reduce licensing costs, given that most software libraries and tools are either free to use or come at a lower cost compared with proprietary alternatives. From a public sector perspective, the adoption of OSS can result in substantial cost savings. Furthermore, OSS adoption brings various benefits to a country across multiple domains, including politics, economy, society, and technology [21].

Examples of Oss Application in AI Software

OSS applications have been extensively utilised in AI software. Table 2 provides a comparison of some OSS applications used in AI software, along with their features.

Table 2: Features offered by certain OSS

OSS	Features
ClearML	• Offers a hassle-free ecosystem for managing experiments • Schedule tasks based on priority and resources • Can run Bayesian hyperparameter optimisation effortlessly
H20.ai	• Can integrate with Hadoop and Spark for big data analysis AI modelling • Tons of AI libraries, unsupervised and supervised learning included • Has data ingestion support across different sources with multiple formats available.
MyCroft.ai	• Comes with modular architecture and can replace its internal components • Instant compile, converting natural language processing (NLP) into machine-readable data through Mozilla's common voice partnership
OpenCV	• Contains a vast number of library tools (decision tree learning, K-nearest, artificial neural network, etc.) • Multiple platforms computability (e.g., Android, iOS, Windows MacOS, Linux, etc.)
PyTorch	• Supports algorithms for computer vision and NLP • Supports end-to-end mobile application development, from Python to iOS or Android • Has performance optimization and training distribution.
TensorFlow	• Supported by multi-language programming, including JavaScript • Compatible across multiple TensorFlow versions of trained AI or machine learning • Has various applications, such as predictive analysis, classification of objects and conversational AI

In conclusion, the implementation of OSS in AI software has resulted in the development of many tools and libraries. These resources will significantly contribute to global AI development.

CURRENT ISSUES AND CHALLENGES

Licensing

While OSS is free to use, certain software companies exercise control over the licensing of their software. This approach, known as dual licensing or general public licensing, has gained popularity as companies derive economic benefits while also providing advantages to the community. However, some members of the OSS community argue that dual licensing does not serve the public interest due to licensing limitations. These limitations hinder the full utilisation of OSS capabilities [22].

Data Quality and Quantity

OSS faces challenges concerning the quality and quantity of data. Despite being developed by the community, OSS does not always offer a completely transparent and high-quality dataset. Transparency in OSS requires the consideration of all features. Additionally, OSS AI software may be implemented on "closed data", meaning it is not entirely transparent, and the public cannot access the data. Furthermore, OSS relies on continuous development from volunteers, and there is a risk that developers may abandon the project, leading to its discontinuation [21].

RELATED WORK

Several researchers have used OSS in their research related to AI. The table below presents the findings of these papers.

Table 3: Some related works by other researchers

Authors	Findings
[23]	The authors evaluate the use of TensorFlow, an OSS, as an AI software for classifying single-cell images from cervical cytology samples. The algorithm achieved an accuracy of 80% for conventional Pap smears and 86% for ThinPrep samples.
[24]	These authors discuss the importance of open-source platforms and frameworks for AI and machine learning. The paper highlights the significance of diversity in these fields to ensure fairness in data-driven decisions made by AI and machine learning algorithms.
[25]	This paper discusses the potential of using OSS for AI and the associated challenges. It proposes seven recommendations to overcome these challenges and promote the use of open-source AI, including defining an OSS AI assessment policy, performing a case study on OSS AI tools, and promoting the use of open standards.
[26]	This paper explores the benefits of using OSS in machine learning. OSS can help increase the reproducibility of research, facilitate fair comparisons between algorithms, and allow for quicker detection and correction of bugs. It also provides access to bug fixes and extensions from external sources, more efficient algorithms, and leverage of existing resources to aid new research.
[27]	This research paper provides a comprehensive literature review on the advantages and disadvantages of AI and machine learning, aiming to assist information professionals in academic research.
[28]	The research addresses comprehension and retention issues in students through AI and ML algorithms, highlighting the potential benefits and advancements associated with these research efforts.

CONCLUSION

In summary, OSS holds immense potential to shape the AI industry positively. Appropriate regulations should be put in place to harness the benefits of OSS effectively. Governments should ensure open access to high-quality data for public institutions, fostering innovation in AI technology within the OSS community. Furthermore, collaboration between the government and large corporations to integrate their technologies into a cohesive ecosystem is crucial. This integration can bolster the economy, fortify national security against threats, and generate new job opportunities within the AI industry. However, regulations overseeing OSS development are essential to prevent issues like data breaches and complications that could negatively impact society.

CONFLICTS OF INTEREST

The authors declare that they have no conflict of interest.

REFERENCE

[1] Ehioghae, E., & Idowu, S. (2021). Open-source software in emerging technologies for economic growth. *ITEGAM- Journal of Engineering and Technology for Industrial Applications* (ITEGAM-JETIA), *7*(27), 63-69. https://doi.org/10.5935/jetia.v7i27.730

[2] Höst, M., & Oručević-Alagić, A. (2011). A systematic review of research on open source software in commercial software product development. *Information and Software Technology*, *53*(6), 616-624. https://doi.org/10.1016/j.infsof.2010.12.009

[3] Xing, M. (2011). Open source software and R&D competition in software industries. *Procedia Engineering*, *15*, 1377–1381. https://doi.org/10.1016/j.proeng.2011.08.255

[4] Goedeken, E. A. (2015). History with an impact: The most cited articles in the journal of library history and its successors over the past fifty years. *Information and Culture*, *50*(2), 285-314. University of Texas Press. http://doi.org/10.7560/IC50301

[5] Kavanagh, P. (2004). *I Open Source Software: Definitions and History* [Online]. Available in http://www.gnu.org/

[6] Dell'Era, C., Di Minin, A., Ferrigno, G., Frattini, F., Landoni, P., & Verganti, R. (2020). Value capture in open innovation processes with radical circles: A qualitative analysis of firms' collaborations with Slow Food, Memphis, and Free Software Foundation. *Technological Forecasting and Social Change*, *158*, 120128. https://doi.org/10.1016/j.techfore.2020.120128

[7] M. Gael Le, ."2.2 The GNU project," 2005. Accessed: Dec. 13, 2023. [Online]. Available: https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=07d2e5208664d740ce649f29cf6ff68186e9a6f5[8] Williams, S. (2010). *Free as in freedom (2.0): Richard Stallman and the free software revolution*. Free Software Foundation.

[9] Von Krogh, G., & Von Hippel, E. (2006). The promise of research on open source software. *Management Science*, *52*(7), 975-983. https://doi.org/10.1287/mnsc.1060.0560

[10] Nelson, M., Sen, R., & Subramaniam, C. (2006). Understanding open source software: A research classification framework. *Communications of the Association for Information Systems*, *17*, 266-288. https://doi.org/10.17705/1CAIS.01712

[11] Zhang, C., & Lu, Y. (2021). Study on artificial intelligence: The state of the art and future prospects. *Journal of Industrial Information Integration*, *23*, 100224. https://doi.org/10.1016/j.jii.2021.100224

[12] Nathan, T. S. S., Kavithambika, S., & Palanichamy, N. V. (2023). Artificial intelligence in agriculture. In An anthology of agricultural research (pp.126-134). https://www.researchgate.net/publication/368543766_Artificial_Intelligence_in_Agriculture

[13] Ku Chhaya, K. A., & Sarode, R. D. (2020). Advantages and disadvantages of artificial intelligence and machine learning: A literature review. *International Journal of Library & Information Science* (IJLIS), *9*(1), 30-36. https://iaeme.com/Home/issue/IJLIS?Volume=9&Issue=1

[14] Bhattacharya, S., Czejdo, B., Agrawal, R., Erdemir, E., & Gokaraju, B. (2018). Open source platforms and frameworks for artificial intelligence and machine learning. *Conference Proceedings - IEEE SoutheastCon*, 2018, St. Petersburg, FL, USA (pp. 1-4). https://doi.org/10.1109/SECON.2018.8479098

[15] R. Ghioni, M. Taddeo, and L. Floridi, "Electronic copy available," 2022. Accessed: Dec. 13, 2023. [Online]. Available: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4272245[16] Bahamdain, S. S. (2015). Open source software (OSS) quality assurance: A survey paper. *Procedia Computer Science*, *56*, 459-464. Elsevier. https://doi.org/10.1016/j.procs.2015.07.236

[17] Zhonghua, Q., & Brocklehurst, M. (2003). What will it take for China to become a competitive force in offshore outsourcing? An analysis of the role of transaction costs in supplier selection. *Journal of Information Technology*, *18*(1), 53-67. https://doi.org/10.1080/0268396031000077459

[18] Martikkala, A., David, J., Lobov, A., Lanz, M., & Ituarte, I. F. (2021). Trends for low-cost and open-source IoT solutions development for industry 4.0. *Procedia Manufacturing*, *55*, 298-305. Elsevier. https://doi.org/10.1016/j.promfg.2021.10.042

[19] Khandelwal, A. (2020). Impact of open source software in research. http://dx.doi.org/10.46890/SL.2020.v01i08.002

[20] Dayaratna, A. (2019). How open source is the key to innovation, productivity, collaboration, and transparency within the digital enterprise. IDC. https://assets.ctfassets.net/wfutmusr1t3h/2ZmFv5YNb3pyErnXcCaxrn/7f828ee74aa920eb9adaca267093d9f2/_downloads_IDC.Report.-.How.open.source.is.the.key.to.innovation.pdf

[21] Theben, A., & Gunderson, L. (2021). Challenges and limits of an open source approach to artificial intelligence policy department for economic, scientific and quality of life policies directorate-general for internal policies [Online]. Available: http://www.europarl.europa.eu/supporting-analyses

[22] Välimäki, M. (2003). Dual licensing in open source software industry affiliation: Researcher at the Helsinki Institute for Information Technology [Online]. Available: http://ssrn.com/abstract=1261644Electroniccopyavailableat:https://ssrn.com/abstract=1261644Electroniccopyavailableat:http://ssrn.com/abstract=1261644Electroniccopyavailableat:https://ssrn.com/abstract=1261644

[23] Van Smaalen, K., Kutob, L., Griffith, C., Mosunjac, M., & Hanley, K. (2019). Distribution of surgical specimen results for Paps PST63 practice habits and adherence to LAST recommendations to resolve discrepancies between cervical biopsies and LEEP/Cone Specimens [Online]. Available: http://mde-lab.aegean.gr.

[24] S. Bhattacharya, B. Czejdo, E. Erdemir, R. Agrawal, and B. Gokaraju, "Open Source Platforms and Frameworks for Artificial Intelligence and Machine Learning," 2018. Accessed: Dec. 13, 2023. [Online]. Available: https://www.researchgate.net/publication/328179850_Open_Source_Platforms_and_Frameworks_for_Artificial_Intelligence_and_Machine_Learning[25] Theben, A., & Gunderson, L. (2021). Challenges and limits of an open source approach to

artificial intelligence policy department for economic, scientific and quality of life policies directorate-general for internal policies [Online]. Available in http://www.europarl.europa.eu/supporting-analyses

[26] Ehioghae E., & Idowu, S. (2021). Open-source software in emerging technologies for economic growth. *ITEGAM- Journal of Engineering and Technology for Industrial Applications* (ITEGAM-JETIA), *7*(27), 63-69. https://doi.org/10.5935/jetia.v7i27.730