AIP Conference Proceedings

AIP Publishing

Volume 3077

Proceedings of the Transdisciplinary Symposium on Engineering and Technology (TSET) 2022

Development of Digital and Green Technology on Post Pandemic Era

Yogyakarta, Indonesia • 21 September 2022

Editors • Ade Gafar Abdullah, Desi Ramayanti, Henri Septanto and Yohanes Galih Adhiyoga

Available Online: pubs.aip.org/aip/acp

RESEARCH ARTICLE | JULY 12 2024

Preface: Proceedings of the Transdisciplinary Symposium on Engineering and Technology (TSET) 2022 FREE

AIP Conf. Proc. 3077, 010001 (2024)
https://doi.org/10.1063/12.0024850

PREFACE: Proceedings of the Transdisciplinary Symposium on Engineering and Technology (TSET) 2022

"Development of Digital and Green Technology on Post Pandemic Era"

It is with great pleasure to welcome you to Transdisciplinary Symposium on Engineering and Technology (TSET) 2022 hosted by Universitas Dian Nusantara on September 21, 2022. The event aims to a venue for engineers, researchers, scholars, and policy makers to explore the challenges and opportunities from the post pandemic era on civil engineering, mechanical engineering, electrical engineering and computer science. For civil engineers, they will play a significant part in the recovery since design and construction services will be needed in the future, and they need to develop new construction methods, materials, and technologies in order to build a sustainable and resilient infrastructure. For engineers, they need to start thinking about the long-term change of their operations and adapt to the "new normal" that has emerged because of the epidemic. We welcome all parties to share their research and thoughts in the symposium.

Participants of the symposium were invited to submit their papers and disseminate them through oral presentation covering such scope as civil engineering, mechanical engineering, electrical engineering and computer science. To enrich the discussion under the theme of "Development of Digital and Green Technology on Post Pandemic Era", we invited speakers with reputable expertise, namely Prof. Josaphat Tetuko Sri Sumantyo, Ph.D. from Chiba University, Japan; Prof. Dr. rer. nat. Evvy Kartini, M.Sc. from National Nuclear Energy Agency of Indonesia; Prof. Dr. Ir. Bambang Sugiarto,, M.Eng. from Universitas Indonesia, Indonesia; and Sulfikar Amir, Ph.D. from Nanyang Technological University, Singapore. In addition to presenting their research results, the participants of the symposium were also encouraged to submit their papers to be proposed for publication to American Institute of Physics (AIP), one of the world's top publishers as conference proceedings. There were 125 manuscripts submitted to the committee comprising 99 papers of Biology, Chemistry, Computer Science and Technology, and Engineering.

Finally, on behalf of the editors of TSET 2022, I would like to extend my most sincere gratitude to the organizing committee, co-hosting institutions, and most importantly, participants, speakers, presenters, and authors of the symposium. I do hope the proceedings bring significant contribution, particularly to the field of advances of sustainable engineering. I look forward to seeing you all at the upcoming symposium.

The Editors,
Ade Gafar Abdullah
Desi Ramayanti
Henri Septanto
Yohanes Galih Adhiyoga

Proceedings of the Transdisciplinary Symposium on Engineering and Technology (TSET) 2022
AIP Conf. Proc. 3077, 010001-1–010001-1; https://doi.org/10.1063/12.0024850

RESEARCH ARTICLE | JULY 12 2024

Committees: Proceedings of the Transdisciplinary Symposium on Engineering and Technology (TSET) 2022 FREE

AIP Conf. Proc. 3077, 010002 (2024)
https://doi.org/10.1063/12.0026137

TSET 2022 COMMITTEE

Advisory Board

Prof. Josaphat Tetuko Sri Sumantyo
Prof. Sulfikar Amir
Prof. E.S. Margianti
Prof. Bambang Sugiarto
Prof. Raihan
Prof. Moermahadi Soerja Djanegara
Prof. Evvy Kartini
Prof. Memen Kustiawan
Prof. Jony Oktavianto
Prof. Intiyas Utami
Dr. Awaludin Martin

Scientific Committee

Prof. Ade Gafar Abdullah
Prof. Suharyadi
Dr. M. Hasannudin Toyieb
Dr. Dewi Anggraini
Magito

Organizing Committee

Conference Chair:

Ir. Margono Sugeng, M.Sc.

Co-Conference Chair:

Ir. Komarudin, ST., MT

Members:

Assoc. Prof. Isma Widiaty
Dr. Yohanes Galih Adhiyoga
Dr. Ari Purwanti
Dr. Didin Hikmah Perkasa
Kornelia Johana, S.I.Kom., M.I.Kom

Proceedings of the Transdisciplinary Symposium on Engineering and Technology (TSET) 2022
AIP Conf. Proc. 3077, 010002-1–010002-1; https://doi.org/10.1063/12.0026137

RESEARCH ARTICLE | JULY 12 2024

Improving of anticancer compound identification model for medicinal plant's LCMS data

Iwan Binanto

AIP Conf. Proc. 3077, 040017 (2024)
https://doi.org/10.1063/5.0201202

Improving of Anticancer Compound Identification Model for Medicinal Plant's LCMS Data

Iwan Binanto[1, a)]

[1] *Informatics Department, Sanata Dharma University, Yogyakarta, Indonesia*

[a)] Corresponding author: iwan@usd.ac.id

Abstract. The development of an identification tool for anticancer chemical compounds on LCMS data of Medicinal Plant has been successfully carried out in previous studies. It needs further improvement especially in software architecture caused by utilizing Webscraping technique to get data. The utilized of this technique caused in frequent disconnections. Because it is considered as a Denial of Service attack. This paper provides a design to improve the existing architecture so that identification devices can run more smoothly and quickly without being suspected of being a Denial of Service attack. For this reason, a new model was developed from the previous model. The new model is 96.1% faster than the old one for labeling data. This provides huge efficiencies.

INTRODUCTION

Liquid Chromatography and Mass Spectrometry (LC-MS) are often used to identify chemical compounds in medicinal plants, because of their high sensitivity and wide coverage [1,2]. LC-MS allows the physical separation of thousands of metabolites thereby providing more comprehensive information about the metabolites being tested [3–6]. Assays using LC-MS are very important in plants, because plant biochemistry is very rich, it has many semi-polar compounds, including major secondary metabolite groups, which can be separated and detected by LC-MS [7].

It is difficult in Liquid Chromatography to ascertain a particular chemical in the peak, even if the sample contains only one chemical. For this reason, it is necessary to add Mass Spectrometry which will provide information about the mass of all chemicals at their peak, so that it can be used to identify them [8].

Mass Spectrometry is based on the analysis of ions moving through a vacuum. The result is a mass spectrum, which provides valuable information about the molecular weight, structure, identity, amount and purity of the sample [8,9]. In short, Mass Spectrometry is used to measure the mass to charge ratio of charged particles which is called the mass to charge ratio (m/z) [8].

The combination of Liquid Chromatography with Mass Spectrometry (LC-MS) allows for a more definitive identification and facilitates the quantitative determination of compounds [10]. The result of LC-MS is a scattered 3D signal with intensity, m/z, and retention time for each detected feature (peak mass). These 3D signals are then converted into 2D signals for easy analysis and interpretation. Usually m/z with intensity is used, although it is possible to use m/z with retention time [11].

The LC-MS signal provides information regarding the mass-charge ratio (m/z) of the ionized molecule or molecular fragment to the retention time of the detected ion and must then be mapped to the metabolite to understand the resulting biochemical processes. This mapping is still in the form of interpretation and complex because it requires efficient and accurate identification.

Proceedings of the Transdisciplinary Symposium on Engineering and Technology (TSET) 2022
AIP Conf. Proc. 3077, 040017-1–040017-8; https://doi.org/10.1063/5.0201202

Southeast Asian countries including India and China have been using this traditional plant for years for alternative cancer therapies. One of them is Typhonium Flagelliforme which is a potential health supplement to cure breast, lung, rectal, liver, prostate, pancreatic and cervical cancers as well as leukemia [12–16].

Many studies on medicinal plant especially Typhonium Flagelliforme have been carried out, including those by Sianipar et al [17–23]. In several studies, they tested with LC-MS. This LC-MS data is then interpreted manually to get the names of its chemical compounds.

Processing and analyzing the LC-MS data manually takes a lot of time because it requires extensive processing and involves a lot of data [4,24–27]. So, it takes a lot of time and effort [1,28,29]. Here it is necessary to use computations and algorithms to simplify and speed up the analysis and interpretation of LC-MS data to extract existing information [11].

In previous studies, a tool for detecting chemical compounds in Typhonium Flagelliforme has been developed [30] by utilizing webscraping techniques to gathering data [31].

Webscraping is a method for retrieving, preparing and scraping required content which can be structured or unstructured data from web pages. Usually this data is in the form of patterned data, especially lists or tables [32,33], although it is not limited to such data. In short, webscraping is web mining [34,35].

Utilized this technique in previous studies still has drawbacks, which are frequent disconnections and unsatisfactory in terms of time [31]. In this paper, we will discuss this problem and possible solution.

Liquid Chromatography is a universal technique used for the separation of compounds from mixtures [36] that safely separates a wide range of organic compounds, from small molecules of drug metabolites to peptides and proteins [37].

Webscraping can be thought of as an automation of manual copy-paste jobs. This work is done in an organized and automated manner, by a virtual computer agent. This agent will follow every link from the web page, which actually does the same job as humans normally do when interacting with web pages. These agents can follow links (by providing HTTP GET requests) and submit forms (via HTTP POST), as well as browse many different web pages. This is done at super-fast computer speeds than humans do [35]. Therefore, sending HTTP requests at computer speed, can be a problem because the server will assume that someone is trying a Denial-of-Service attack, due to the large number of requests in a very short time [38].

This research will improve the previous model so that connections do not disconnect frequently and speed up data labeling with webscraping.

METHODS

Analysis

The problems that arise in previous studies are frequent disconnections and unsatisfactory in terms of time. Frequent disconnections can be caused by a bad network or indeed being disconnected by the server because it is considered an attack namely Denial-of-Service attack.

It because the way of webscraping works is to provide HTTP requests at computer speed. This arise a problem because the server will assume that someone attempted a Denial-of-Service attack, due to the large number of requests in a very short time. This caused to be unsatisfactory in terms of time because it becomes slow and repeatedly disconnects.

In previous studies, we developed a tool based on the model as shown in Figure 1. It can be seen that there is a direct connection to an online site which is an online database, namely massbank.jp [39] and NPACT [40]. This direct connection uses webscraping technique to get the data provided by massbank.jp and NPACT as shown in Figure 1. This is the main problem as circled in black in Figure 2. It can be considered as Denial-of-Service attack, which often lead disconnections.

An initial solution to overcome the assumption of a DoS attack is utilizing delays for every requests that made. But it doesn't solve the frequent disconnections perfectly. So, the existing model needs to be modified to solve the problem.

Design The Model

Analysis that has been carried out provides idea for modifications to the existing model. This section will discuss the design of the new modified model.

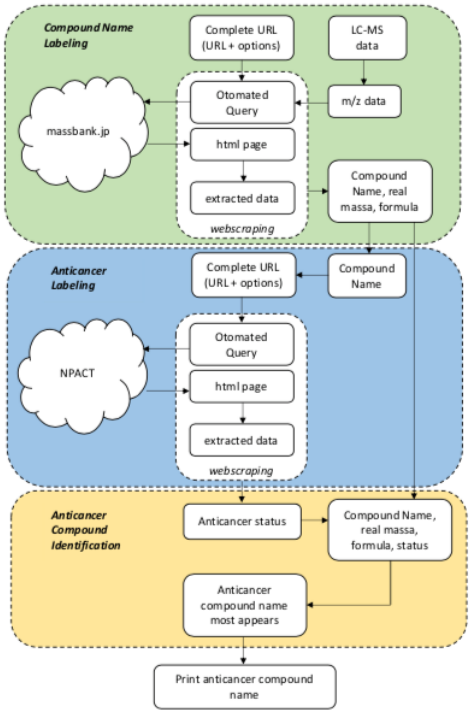

FIGURE 1. Previous model [30].

Main problem in here is the direct connection to the online database as circled in black in Figure 2. So, the solution that will be used is to build a mirror database, so it does not connect directly to the online database.

Massbank provides a documentation for build mirror database [41]. This documentation provides an overview of how the mirror database will be set up.

The previous model was developed by placing a mirror database where the problem is circled in black. the developed model as seen as Figure 3.

This model places the mirror database as an intermediary to the online database server, and acts like proxy server. So, the model will require synchronization to keep the mirror database up to date. Synchronization with the actual server, will be done manually in the meantime. The way to synchronize manually is to download the SQL query that has been provided by MassBank on its github site.

Design System Architecture

After designing new model, the system architecture is also designed to be easier to implement. Designing the system architecture means designing the physical connectivity of the computer. Then in this section the physical architecture of the network that will be used will be designed. This design as shown in Figure 4.

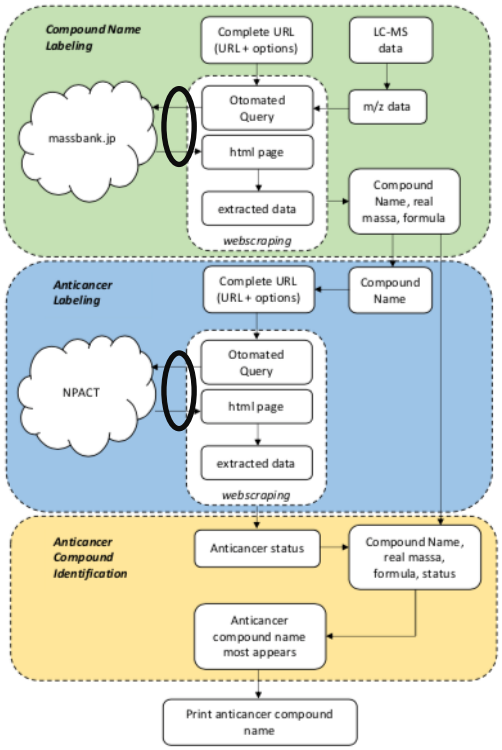

FIGURE 2. Main problem spot in model.

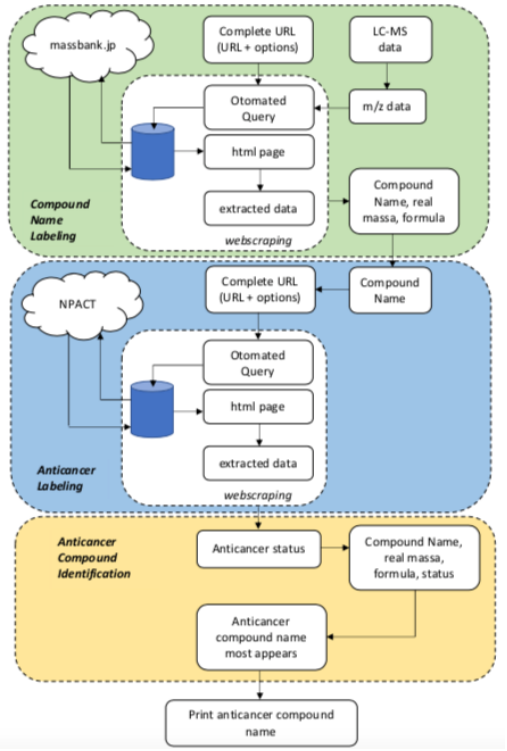

FIGURE 3. Developed model.

In this architecture, two computers are required, one acting as a client and the other as a server. They are connected via a wifi intranet. Connection to the internet is only used for necessary updates. The server will contain a mirror of the MassBank website and its database so that clients only access it.

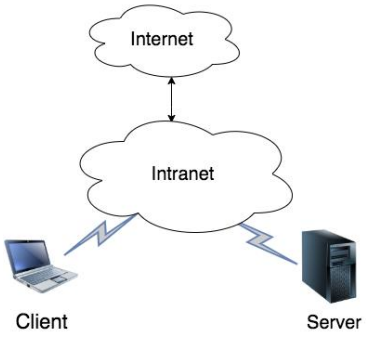

Figure 4. Design system architecture 1.

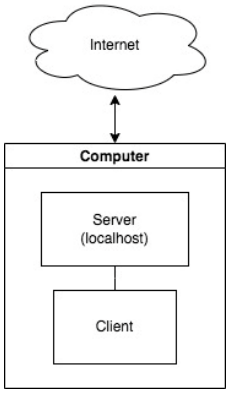

Figure 5. Design system architecture 2.

The system architecture design as shown in Figure 4 can be simplified by only accessing the localhost server as depicted in Figure 5. So, only one computer is needed as it functions as both server and client. Connection to the internet is only used for necessary updates. With this architecture, no connection to other computers is required so there is no physical data communication between computers. In theory, this architecture is getting faster because there is no transmission medium.

RESULTS AND DISCUSSIONS

The analysis of the previous model provides an overview of the existing shortcomings. Therefore, a new model was designed that corrects the existing deficiencies. In this case by build a new model using a mirror database. In implementation, system architecture 1 implemented as shown in the Figure 6 and computer technical specifications are shown in Table 1.

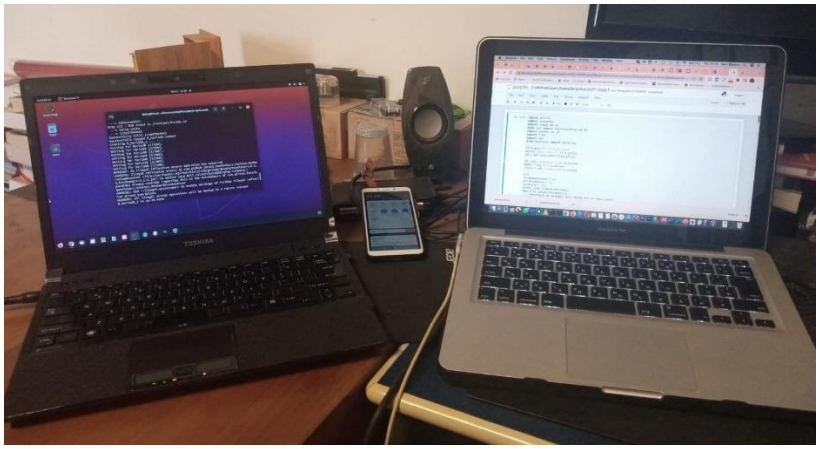

FIGURE 6. Implementation system architecture 1.

In theory, the speed of an intranet is faster than the internet due to the transmission medium and the distance. Likewise, if only use one computer, it will definitely be faster than connecting to the internet or using two computers connected via the intranet.

TABLE 1. Computer technical specifications.

Laptop 1 (server)		Laptop 2 (client)	
Processor	: i3	Processor	: i7
RAM	: 8 GB	RAM	: 16 GB
Storage	: 256 GB	Storage	: 256 GB
Sistem operasi	: Linux UBUNTU	Sistem Operasi	: Mac OS X

In previous studies [31], the most important thing is labeling data which takes a long time, which is about 3 hours for 10,000 records. This is due to the connection to internet and the strategy so that it is not considered a DoS attack. Therefore, only data labeling utilizing MassBank is discussed in this paper.

MassBank server was installed to local server using source from github as mentioned in the provided documentation [41]. During installation, there are several configuration files that must be changed, especially for the username and password for database engine which is MariaDB. It is not easy for an ordinary user and/or novice administrator.

Utilizing this architecture is greatly speeds up the data labeling process. In the experiment with two computers as client and server as system architecture 1, it only took about 7 minutes for 10,000 records. This means speeding up the process up to 96.1%. This is a huge saving time for labeling data. The result of a comparison of the previous model with the new model is shown in Table 2.

TABLE 2. Previous model and new model time comparison.

Previous Model (per 10,000 records)	New Model (per 10,000 records)
3 hours (180 minutes)	7 minutes

This research succeeded improving the previous model by modifying the system architecture from internet-based to intranet-based. This eliminates frequent diconnections and perceived Dos threats from servers.

CONCLUSIONS

The design for the development of the model was successfully carried out by analyzing the shortcomings of the previous model. However, this new model requires further experiments to test the model. The implementation of this model is not so easy but it can be proven that this model can be used and speed up the process up to 96.1%.

The benefit of this research is that it contributes to a new method or model or system architecture for the identification of chemical compounds in medicinal plants, especially Typhonium Flagelliforme and can provide more accurate and speed up information about chemical compounds contained in medicinal plants, thus helping Biological scientists and Pharmaceutical Scientists in utilization of medicinal plants especially Typhonium Flagelliforme.

ACKNOWLEDGMENTS

Researcher express many thanks to Dr. Nesti F. Sianipar, SP, M.Sc. and Ms. Khoirunnisa Assidqi, S.Pi., M.Sc. which has supported in previous research and LC-MS data.

REFERENCES

1. A. Vaniya and O. Fiehn, "Using fragmentation trees and mass spectral trees for identifying unknown compounds in metabolomics," TrAC - Trends Anal. Chem., vol. **69**, pp. 52–61, (2015).
2. B. Zhou, "Computational Analysis of LC-MS/MS Data For Metabolite Identification," Virginia Polytechnic Institute and State University, (2012).
3. I. Blaženović, T. Kind, J. Ji, and O. Fiehn, "Software tools and approaches for compound identification of LC-MS/MS data in metabolomics," Metabolites, vol. **8**, no. 2, (2018).
4. M. Katajamaa and M. Orešič, "Data processing for mass spectrometry-based metabolomics," J. Chromatogr. A, vol. **1158**, no. 1–2, pp. 318–328, (2007).
5. E. R. Perez, J. A. Knapp, C. K. Horn, S. L. Stillman, J. E. Evans, and D. P. Arfsten, "Comparison of LC-MS-MS and GC-MS analysis of benzodiazepine compounds included in the drug demand reduction urinalysis program," J. Anal. Toxicol., vol. **40**, no. 3, pp. 201–207, (2016).
6. B. Rushing, A. Wooten, M. Shawky, and M. I. Selim, "Comparison of LC – MS and GC – MS for the Analysis of Pharmaceuticals and Personal Care Products in Surface Water and Treated Wastewaters," Curr. Trends Mass Spectrom., no. July, pp. 14–17, (2016).
7. R. C. H. De Vos, S. Moco, A. Lommen, J. J. B. Keurentjes, R. J. Bino, and R. D. Hall, "Untargeted large-scale plant metabolomics using liquid chromatography coupled to mass spectrometry," Nat. Protoc., vol. 2, no. 4, pp. 778–791, (2007).
8. P. R. Kumar, S. R. Dinesh, and R. Rini, "Lcms- A Review and a Recent Update," J. Pharm. Pharm. Sci., vol. **5**, no. 5, pp. 377–391, (2016).
9. Agilent, "Agilent 6400 Series Triple Quadrupole LC / MS System Concepts Guide," Agil. Technol., (2012).
10. R. E. Ardrey, Liquid Chromatography –Mass Spectrometry: An Introduction. (2003).
11. F. Fernández-Albert, "Machine Learning Methods for the Analysis of Liquid Chromatography-Mass Spectrometry datasets in Metabolomics," Universitat Polit` Ecnica De Catalunya, (2014).
12. A. Bustamam et al., "In vitro ultramorphological assessment of apoptosis on CEMss induced by linoleic acid-rich fraction from Typhonium flagelliforme tuber," Evidence-based Complement. Altern. Med., vol. **2011**, (2011).
13. C. Y. Choo, K. L. Chan, K. Takeya, and H. Itokawa, "Cytotoxic activity of Typhonium flagelliforme (Araceae)," Phyther. Res., vol. **15**, no. 3, pp. 260–262, (2001).
14. S. T. Gopukumar, P. Alexander, T. K. S. Fathima, R. Jeniffa, and P. K. Praseetha, "Typhonium Divaricatum : a Promising Endangered Anti-Cancer," Med. Plants Their Conserv. Manag., no. July, pp. 114–119, (2015).
15. S. Mohan et al., "Typhonium flagelliforme inhibits the proliferation of murine leukemia WEHI-3 cells in vitro and induces apoptosis in vivo," Leuk. Res., vol. **34**, no. 11, pp. 1483–1492, (2010).
16. M. Singh, D. Kumar, D. Sharma, and G. Singh, "Typhonium Flagelliforme: a Multipurpose Plant," Int. Res. J. Pharm., vol. **4**, no. 3, pp. 45–48, (2013).

17. D. Laurent, N. F. Sianipar, Chelen, Listiarini, and A. Wantho, "Analysis of Genetic Diversity of Indonesia Rodent Tuber (Typhonium flagelliforme Lodd.) Cultivars Based on RAPD Marker)," in *The 3rd International Conference on Biological Science 2013 (The 3rd ICBS-2013), 2015*, vol. 2, pp. 139–145, (2015).
18. N. F. Sianipar, W. Maarisit, and A. Valencia, "Toxic Activities Of Hexane Extract And Column Chromatography Fractions Of Rudent Tuber Plant (Typhonium Flagelliforme Lodd.) On Artemia Salina," Indones. J. Agric. Sci., vol. **14**, no. 1, pp. 1–6, (2013).
19. N. F. Sianipar, R. Purnamaningsih, I. Darwati, D. Laurent, and Chelen, "The Effects Of Gamma Irradiation And Somaclonal Variation On Morphology Variation Of Mutant Rodent Tuber (Typhonium flagelliforme Lodd.) Lines," in *The 3rd International Conference on Biological Science 2013*, vol. 2, pp. 637–645, (2015).
20. N. F. Sianipar and R. Purnamaningsih, "Enhancement of the Contents of Anticancer Bioactive Compounds in Mutant Clones of Rodent Tuber (Typhonium flagelliforme Lodd.) based on GC-MS Analysis," Pertanika J. Trop. Agric. Sci., vol. **41**, no. 1, pp. 305–320, (2018).
21. N. F. Sianipar, A. Wantho, Rustikawati, and W. Maarisit, "The Effects of Gamma Irradiation on Growth Response of Rodent Tuber (Typhonium flagelliforme Lodd.) Mutant in In Vitro Culture," HAYATI J. Biosci., vol. **20**, no. 2, pp. 51–56, (2013).
22. N. F. Sianipar, K. Assidqi, R. Purnamaningsih, and T. Herlina, "in Vitro Cytotoxic Activity of Rodent Tuber Mutant Plant (Typhonium Flagelliforme Lodd.) Against To Mcf-7 Breast Cancer Cell Line," Asian J. Pharm. Clin. Res., vol. **12**, no. 3, pp. 185–189, (2019).
23. N. F. Sianipar, R. Purnamaningsih, D. L. Gumanti, Rosaria, and M. Vidianti, "Analysis of Gamma Irradiated-Third Generation Mutants of Rodent Tuber (Typhonium flagelliforme Lodd.) Based on Morphology, RAPD, and GC-MS Markers," Pertanika J. Trop. Agric. Sci., vol. **40**, no. 1, pp. 185–202, (2017).
24. R. Danielsson, D. Bylund, and K. E. Markides, "Matched filtering with background suppression for improved quality of base peak chromatograms and mass spectra in liquid chromatography-mass spectrometry," Anal. Chim. Acta, vol. **454**, no. 2, pp. 167–184, (2002).
25. R. Tautenhahn, C. Bottcher, and S. Neumann, "Highly sensitive feature detection for high resolution LC/MS," BMC Bioinformatics, vol. **9**, no. 504, pp. 1–16, (2008).
26. F. Fern, R. Llorach, C. Andr, and A. Perera, "Supplementary DATA: An R package to process LC / MS metabolomic data : MAIT (Metabolite Automatic Identification Toolkit)," pp. 1–17, (2016).
27. E. Gorrochategui, J. Jaumot, and R. Tauler, "ROIMCR: A powerful analysis strategy for LC-MS metabolomic datasets," BMC Bioinformatics, vol. **20**, no. 1, pp. 1–17, (2019).
28. P. J. Ulintz, J. Zhu, Z. S. Qin, and P. C. Andrews, "Improved Classification of Mass Spectrometry Database Search Results Using Newer Machine Learning Approaches," Mol. Cell. Proteomics, vol. **5**, no. 3, pp. 497–509, (2006).
29. R. J. M. Weber et al., "Computational tools and workflows in metabolomics: An international survey highlights the opportunity for harmonisation through Galaxy," Metabolomics, vol. **13**, no. 12, pp. 1–5, (2017).
30. I. Binanto, H. L. H. S. Warnars, N. F. Sianipar, and W. Budiharto, "Anticancer Compound Identification Model Of Rodent Tuber's Liquid Chromatography-Mass Spectrometry Data," ICIC Express Lett., vol. 16, no. 1, pp. 9–16, (2022).
31. I. Binanto, H. L. H. S. Warnars, N. F. Sianipar, and W. Budiharto, "Webscraping Data Labeling System On Liquid Chromatography-Mass Spectrometry Of Rodent Tuber For Efficiency Of Supervised Learning Preprocessing," ICIC Express Lett. Part B Appl., vol. **13**, no. 1, pp. 107–114, (2022).
32. N. R. Haddaway, "The use of web-scraping software in searching for grey literature," Grey J., vol. **11**, no. February, pp. 186–190, (2016).
33. M. Shreesha, S. B. Srikara, and R. Manjesh, "A Novel Approach for News Extraction Using Webscraping Technique," in *The 3rd National Conference on Image Processing, Computing, Communication, Networking and Data Analytics (NCICCNDA 2018)*, no. NCICCNDA 2018, pp. 359–362, (2018).
34. R. S. Chaulagain, S. Pandey, S. R. Basnet, and S. Shakya, "Cloud Based Web Scraping for Big Data Applications," in *Proc. - 2nd IEEE Int. Conf. Smart Cloud, SmartCloud 2017*, pp. 138–143, (2017).
35. A. V Saurkar, K. G. Pathare, and S. A. Gode, "An Overview On Web Scraping Techniques And Tools," Int. J. Futur. Revolut. Comput. Sci. Commun. Eng., vol. **4**, no. 4, pp. 363–367, (2018).
36. M. Ludovici, C. Ialongo, and E. Camera, "Principles, current applications, and future perspectives of liquid chromatography-mass spectrometry in clinical chemistry," in Liquid Chromatography: Applications: Second Edition, Second Edi., vol. **2**, Elsevier Inc., pp. 727–751, (2017).
37. Hewlett Packard, Basics of LC/MS. (2001).

38. F. J. A. P. Mattosinho, “Mining Product Opinions and Reviews on the Web,” Technische Universitat Dresden, (2010).
39. H. Horai et al., “MassBank: A public repository for sharing mass spectral data for life sciences,” J. Mass Spectrom., vol. 45, no. 7, pp. 703–714, (2010).
40. M. Mangal, P. Sagar, H. Singh, G. P. S. Raghava, and S. M. Agarwal, “NPACT: Naturally occurring plant-based anti-cancer compound-activity-target database,” Nucleic Acids Res., vol. **41**, no. D1, pp. 1124–1129, (2013).
41. “Installation of MassBank | MassBank-documentation.”.