Prof. Dr. Moechammad Sarosa (Electrical Engineering Department, Politeknik Negeri Malang, Indonesia).

Dr. Eng. Cahya Rahmad (Information Technology Department, Politeknik Negeri Malang, Indonesia)

Dr. Dewi Yanti Liliana (Information Technology Department, Politeknik Negeri Jakarta, Indonesia).

Dr. F. Yudi Limpraptono (Electrical Engineering Department, Institut Teknologi Nasional, Indonesia).

Dr. Isdawimah (Electrical Engineering Department, Politeknik Negeri Jakarta, Indonesia).

Mohammad Noor Hidayat, ST., M.Sc., Ph.D. (Electrical and Electronics Engineering Department, Politeknik Negeri Malang, Indonesia).

Dr. Noor Cholis Basjaruddin (Electrical Engineering Department, Politeknik Negeri Bandung).

PREFACE

We would like to present, with great pleasure, the second issue of Matrix: Jurnal Manajemen Teknologi dan Informatika in Volume 12, 2022. This journal is under the management of Scientific Publication, Research and Community Service Center, Politeknik Negeri Bali and is devoted to cover the field of technology and informatics management including managing the rapid changes in information technology, emerging advances in electrical and electronics and new applications, implications of digital convergence and growth of electronics technology, and project management in electrical, mechanical or civil engineering. The scientific articles published in this edition were written by researchers from Politeknik Masamy International, Universitas Amikom Yogyakarta, Universitas Flores, Universitas Indonesia, and Universitas Bumigora. Articles in this issue cover topics in the field of Information Management and Electrical Engineering, Hoax news identification using machine learning model from online media in Bahasa Indonesia, Delone and McLean Models for measuring the success of Flores University e-learning information system, Business transformation from connectivity to digital: case study "PT. MNO Indonesia", Asah skill: An approach skill-based learning platform with management and strategic planning process and Application of the Blowfish Algorithm in securing patient da-ta in the database. Finally, we would like to thank reviewers for their efforts and hard work in conducting series of review phase thoroughly based on their expertise. It is our hope that the work of the authors in this issue will be a valuable resource for other researchers and will stimulate further research into the vibrant area of technology and information management in specific, and engineering in general.

Politeknik Negeri Bali, 31 July 2022

Editor-in-chief

Gusti Nyoman Ayu Sukerti, S.S., M.Hum.

TABLE OF CONTENT

Hoax news identification using machine learning model from online media in Bahasa Indonesia

Inggrid Yanuar Risca Pratiwi [1], **Anggit Ferdita Nugraha** [2*]

[1] Program Studi Teknologi Komputer, Politeknik Masamy International, Indonesia
[2] Program Studi Teknik Komputer, Fakultas Ilmu Komputer, Universitas Amikom Yogyakarta, Indonesia

*Corresponding Author: *anggitferdita@amikom.ac.id*

Abstract: Information and communication technology that's developing is one of the main triggers of the information explosion today. Nowadays, various news content is not only easy to obtain but also easy to produce through various platforms on the internet, including popular online media, such as blogs and websites. So a lot of news content on blogs and websites that are currently being circulated leads to fake news content (hoaxes) that can mislead the perception and thoughts of the readers. Therefore, it is important to develop a system that can detect the presence of fake news content to minimize the losses caused by the presence of fake news content. In this study, the Naive Bayes algorithm is proposed as a machine learning model that will be used to detect fake news content in Indonesian language online media. As a result, the global accuracy value reached 71% with recall, precision, and F1-Score values as a whole above 70% which indicates that the proposed model can detect fake news content quite well.

Keywords: listing important terms, enabling readers to find the article, in range of 3-6 terms

How to Cite: I. Y. R. Pratiwi and A. F. Nugraha, "Hoax news identification using machine learning model from online media in Bahasa Indonesia," Matrix: Jurnal Manajemen Teknologi dan Informatika, vol. 12, no. 2, pp. 58-67, 2022.

Introduction

The increasingly massive development of information and communication technology makes the world seems borderless [1]–[3]. As a result, information is now easier, faster, and more flexible to be accessed anywhere and anytime regardless of the location and place where the information is located. These developments certainly cannot be separated from the role of the internet which is currently evolving continuously and has even transformed into a major need for humans [4]–[6].

The internet, with its various advantages, is now able to become the main axis as well as an important factor in the development of modern and online-based information media such as social media and publication media such as websites and blogs. Websites and Blogs are currently growing and have become one of the effective information publication media, especially those related to news channels [7]–[9]. In fact, with the development of Websites and Blogs as media for publication at this time, the existence of traditional information media such as radio and television, as well as print media in the form of magazines, newspapers, and tabloids are now gradually being abandoned [10], [11].

Websites and blogs that make it easy for anyone not only to access information easily but also to be able to share information very quickly and freely have resulted in a massive information explosion [12]. The biggest impact caused is a change in mindset and the splitting of one's perception and belief in understanding the information received, this kind of thing, if left unchecked, can be the main trigger for a split between groups due to news and information whose validity is doubtful [8], [9].

It is important for someone to always increase awareness and vigilance when accessing news and information content through websites, blogs, and other online media to always ensure the truth and validity of the news displayed. Support for systems and applications that can assist

someone in the early detection of fake news content is also very much needed to minimize the impact.

Until now, the researchers are still developing the systems and applications for the early detection of fake news content in online media. Various methods are proposed to find the maximum performance in detecting the presence of fake news content on a website page. One of the popular methods used in the research is the machine learning method [13]–[16]. Popular algorithms like Support Vector Machine (SVM) [13]–[15], Logistic Regression [14], [15], and Naive Bayes [5], [8], [13]–[16] are widely used and perform well when it comes to detecting fake news content in various datasets. As shown in research for detecting fake news from social media data like WhatsApp messaging and Facebook posts, the Application of SVM and the Naïve Bayes algorithm resulted in an accuracy of up to 93.5% [13]. Similar findings were discovered in a study conducted on Chile Earthquake 2010 datasets [14]. By using the Naive Bayes algorithm as a machine learning model, the resulting classification performance also shows a good value with an accuracy metric of 89.06% [14]. As a result of these many research, it is clear that the Naive Bayes algorithm has a high potential for detecting false news in a variety of datasets [13]–[16].

Unfortunately, based on the majority of studies, the detection of fake news using naive Bayes is widely applied to English datasets. Therefore, experiments on the usage of the Naïve Bayes classifier as a machine learning model for detecting fake news content in Indonesian language online media will be conducted in this study.

Methodology

The development of a machine learning-based fake news detection system for Indonesian language online media follows the development stages shown in Figure 1.
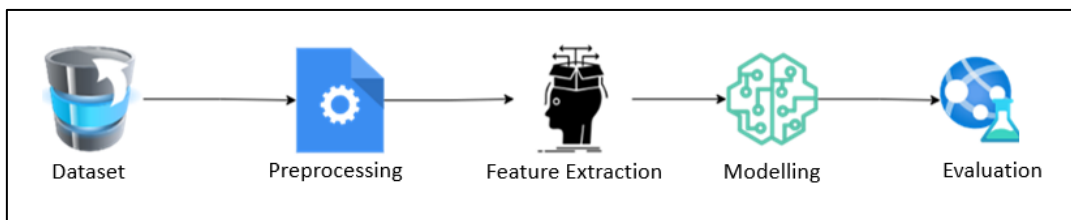
Figure 1. System development stages

In Figure 1, the stages of system development are carried out sequentially starting from the data acquisition process, then followed by the pre-processing stage (pre-process), followed by the feature extraction stage, the next is the modeling stage, and the last stage the performance evaluation. Each process carried out at each stage will be explained in detail and further through the respective sub-chapters.

A.   Dataset

The dataset used in this study is the "Indonesian Hoax News Detection Dataset" which can be downloaded for free through the Mendeley Data Repository website page [8]. The dataset contains 600 Indonesian news articles consisting of 12 topics with 50 news articles for each topic. The dataset is also included in the binary classification where the first class is labeled as Valid News, and the second class is labeled as Fake News (hoax). The labeling process was carried out by 3 annotators manually with the final results taken using a vote tagging mechanism.

To label it as valid news, annotators use references sourced from Indonesian-language online media such as kompas.com, merdeka.com, tribunnews.com, and various other media that have credibility as media whose validity has been tested. While the label for fake news (hoax) will be given by the annotator when the news is obtained from media whose credibility is questionable [8]–[11].

B.   Pre-processing

Pre-processing is the second stage that is carried out after the dataset is acquired. The purpose of doing this pre-processing stage is to understand the characteristics of the data, examine the data, and perform data cleaning so that the data will be ready to use at various

stages of further processing. Some of the processes involved in the pre-processing stage are shown through the flow in Figure 2.

Figure 2. Pre-processing steps

The pre-processing stage in Figure 2 begins with the Case Folding process. This process is one of the common strategies and is widely used in data processing, especially for data in the form of text.
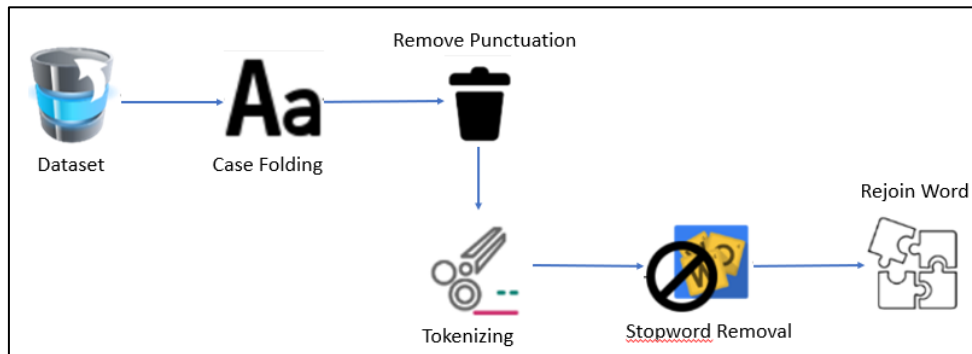
Case Folding is a process to change all the letters in the data, into lowercase, to maintain cosistency and generalize the structure of the text data used [9].

Data that has passed the Case Folding process then continue with the process of deleting punctuation marks. In addition to punctuation, various characters that have no meaning, and have no effect on processing text data will also be deleted and removed through the mechanism of the Remove Punctuation process.

Next is the tokenizing process, wherein in this process, the sentences in the data text will be cut into several words while removing certain characters that are not used in the study based on the presence of spaces in the data. The purpose of doing this tokenizing process is to find out what words are influential and tend to represent data from a certain class [3].

As a result of the tokenizing process, stop word removal will then be carried out to remove general words such as: "yang", "this", "is", "ke", and "di", where these words tend to have a high frequency of occurrence but do not have a specific meaning [8], [9].

The purpose of the stop word removal process is to speed up process performance because it reduces some common words, as well as to improve model performance because the data to be modeled is data that specifically contains important words that represent a certain class [4]–[6], [11]. For the stop word removal process carried out in this study, the researcher used the Indonesian language stop word list from Talla F.Z [17].

Finally, the process carried out from this pre-processing stage is the Rejoin Word process wherein this process has done to combines various words that have been processed through various previous stages so that they return to a unified whole sentence and become the main data that is ready to be processed using a machine model learning in the next processing stage [8], [9].

C.   Feature Extraction
The Feature Extraction stage is the processing stage that is carried out after passing the pre-processing stage. In this study, the feature extraction process was carried out using the n-gram method and the Term Frequency - Inverse Document Frequency (TF-IDF) method. An n-gram is a form of feature extraction that works by breaking sentences into a set of n-word combinations [3], [7].

In this study, the use of the n-gram method is carried out in the form of trigrams (3-grams), where each sentence in the data will be cut by 3 words based on the proximity of the words, and each word resulting from the trigram process has then calculated the weight of each word using Term Frequency – Inverse Document Frequency (TF-IDF). The goal is to find out how big the relationship of a word is to the number of occurrences in a document.

Calculations performed using Term Frequency - Inverse Document Frequency are carried out by combining Term Frequency calculations to determine the frequency of occurrence of words

in a document, as well as performing Inverse Document Frequency calculations to determine the frequency of occurrence of words in documents so that it can be seen how important the word is in a document [7]. The general equation for calculating TF-IDF is as equation:

$$Wdt = TFdt * IDFdt \tag{1}$$

The W symbol in equation (1) shows the amount of weight that the d-th document has, on the occurrence of the t-th word, in the entire data document. While TF shows the number of occurrences of words in a document, and IDF is Inverse Document Frequency which counts the occurrence of the t-th word in all documents.

To obtain the IDF value, the calculation process is carried out using the following equation:

$$IDF = log\, log\, \frac{N}{dft} \tag{2}$$

In quation (2), the symbol N indicates the number of data documents used, while dft indicates the number of documents containing the t-word.

D.   Modeling

The steps for modeling the data are carried out after the feature extraction stage is carried out. At this stage, machine learning algorithms are used to model the data, which makes the system seems to be able to predict whether a news item is included in the category of valid news, or fake news (hoax).

In this study, the machine learning model used is Naïve Bayes. Naïve Bayes is included in one of the popular algorithms used to overcome various problems related to text classification [2], [9], [13], [16]. This is of course inseparable from the workings of the Naïve Bayes Algorithm which uses Probability and Statistics to model and predict the occurrence of a word in certain data class categories [2], [9]. Equation (3) shows the general equation of the Naïve Bayes Algorithm:

$$P(A|B) = \frac{P(B|A)\, P(A)}{P(B)} \tag{3}$$

Where:

$P(A|B)$: Posterior Probability
$P(B|A)$: Likelihood Probability
$P(A)$: Prior Probability
$P(B)$: Evidence Probability

In equation (3), Posterior probability shows the prediction of the probability value based on the occurrence of one event based on information from other events. Likelihood probability describes an opportunity that states the degree of the possible influence of information on the occurrence of another event. The opportunity contains the probability value of the occurrence of an event that has been previously believed and it could be that the event is influenced by other events, lastly, Probability Evidence shows a constant comparison measure based on the probability of event information [16].

E.   Evaluation

The final step is evaluation, which is used to determine the performance of the model proposed in the research. The Confusion Matrix is one of the most commonly used assessment measures for evaluating classification performance [18].

Table 1. Binary classification confusion matrix

Predicted Class	Actual Class	
	Positive	**Negative**
Positive	True Positive [TP]	False Positive [FP]
Negative	False Negative [FN]	True Negative [TN]

As shown in Table 1, the confusion matrix has a square shape that maps each class with other classes on both the actual and the predicted side of the system [19]. In general, there are four (4) parts to the process of mapping data classes in binary classification, namely:

- True Positive
  True Positive (TP) indicates how much of the data predicted by the system is positive, with the actual value also being positive.
- True Negative
  True Negative (TN) shows how much data the system predicts as negative values, and the actual values are also negative.
- False Positive
  False Positive (FP) indicates how much of the data is predicted by the system as positive values, while the actual values are negative.
- False Negative
  False Negative (FN) shows how much of the data is predicted by the system as a negative value, while the actual value is positive.

The confusion matrix will be used to determine the performance indicators of the classification model, such as accuracy, precision, sensitivity, and F1-Score (Harmonic Mean) to determine a machine learning model's success and classification performance [18], [19]. through equation (4), the performance of the classification model based on the accuracy metric

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{4}$$

The accuracy measure shows how effectively the classification performance, as measured by the percentage value provided by the algorithm, predicts properly based on all prediction findings [19].

In addition, the accuracy metric will be used to evaluate the classification model's performance, as can a precision value calculated using the equation (5).

$$precision = \frac{TP}{TP + FP} \tag{5}$$

The precision value is a measure for evaluating the classification performance in terms of the model's dependability in making positive predictions. In other words, by computing the accuracy number, one may determine how effectively the model identifies fake news from the total results projected to be false.

The next metric that can provide information about the classification model's success rate is the sensitivity matrix, which is used to calculate the proportion of data predicted to be positive by the system out of all data with a positive label with the equation (6).

$$sensitivity \ (recall) = \frac{TP}{TP + FN} \tag{6}$$

the F1-Score was used in this study using equation (7) to assess the proportion of the calculated precision and sensitivity values [19].

$$F1\ Score = \frac{2*Precision*recall}{precision+recall} \tag{7}$$

Results and Discussions

Research related to the analysis and detection of fake news content in Indonesian language online media has been carried out according to the methodology described in the previous section. This research was conducted with an analytical approach using a classification technique to determine a news story in the Indonesian language or online media is included in the category of valid news or fake news (hoax). The dataset used in this study was carried out using the "Indonesian Hoax News Detection Dataset" [8], with the data form as shown using the Table. 2. Some examples of the Indonesian Hoax News Detection Dataset.

Table 2. Examples of the Indonesian hoax news detection dataset

No.	News	Tagging
1	Jakarta, Di jejaring sosial, banyak beredar informasi yang menyebut lele sebagai ikan paling jorok. Dalam sesuap daging ikan lele, terkandung 3000 sel kanker.	Valid
2	Bahaya Mengkonsumsi Ikan Lele Yang Mengandung Sel Kanker - Lele adalah sejenis ikan yang hidup di air tawar. Ikan lele mudah dikenali karena tubuhnya yang licin, agak pipih memanjang, ...	Hoax
3	Akhir-akhir ini kita sering mendapat broadcast informasi via BBM tentang Cara pertolongan pertama pada penderita Stroke, bunyi pesan broadcast tersebut kira-kira sebagai berikut: orang yang kena STROKE mendadak ...	Hoax
4	Mudah melengkungnya casing iPhone 6 Plus menjadi perbincangan hangat belakangan ini. Dan itu bukan hisapan jempol semata. Sudah ada pengguna iPhone 6 Plus yang mengalami kejadian tersebut. ...	Valid
5	Sungguh para oknum Konsultan Jenderal Republik Indonesia (KJRI) di Davao, Filipina ini keblinger. Mereka tega membakar dadak merak Reog Ponorogo beserta gamelan pengiringnya karena dianggap barang mengandung berhala. Tindakan KJRI itu jelas-jelas telah menghancurkan warisan budaya luhur nenek moyang kita dan tidak tahu tentang sejarah Reog Ponorogo yang sudah mendunia dan menjadi aset budaya Bangsa Indonesia di UNESCO. ...	Hoax

Table 2 shows some examples of datasets in the Indonesian Hoax News Detection Dataset. In the dataset, there are 2 main columns, namely news, and tagging.

The news column contains various articles that have been taken from various online news sources in Indonesian according to the topic that is the focus of the preparation of the dataset, while the tagging column is a label given to news articles by the vote tagging process carried out by the annotator during the data labeling process.

From the dataset that has been acquired, the next step that needs to be done is to examine the dataset to find out important information that can support the analysis process before entering the next processing stage.

Some important things such as information from the distribution of data classes are mandatory things that need to be known so that the analysis process runs more optimally.
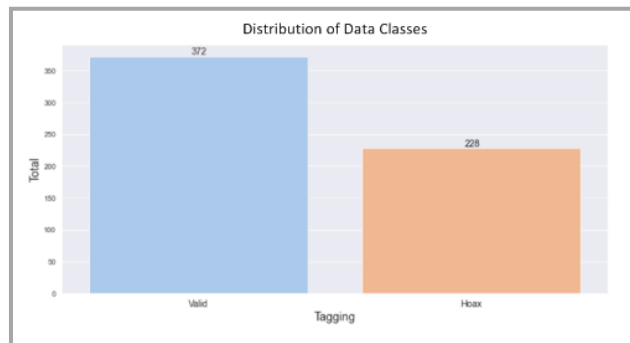
Figure 3. Distribution of data classes

Figure 3 shows that the distribution of data classes between fake news (hoax) and valid news has unbalanced characteristics. For news with valid data class, it has a larger amount of 372 data, while data with a label as fake news (hoax) is 228 data.

Next, the process will be carried out at the pre-processing stage, where the dataset will be processed using a case-folding mechanism to generalize the structure and maintain data consistency. The following is an example of news data that has gone through the case folding process.

Figure 4. Case folding process

Figure 4 shows the news data before and after the case folding process is carried out. For news data that have not gone through the case folding process, it can be seen that the arrangement of letters in news sentences still uses a combination of uppercase and lowercase letters, while for news data that has gone through the case folding process, it can be seen that there are similarities in the structure of letters for each sentence using lowercase letters.

After going through the case folding process to make the structure of the letters the same, the next step is to carry out the remove punctuation process to remove all punctuation marks contained in news sentences. The results of the removal punctuation process are shown in Figure 5.

Figure 5. Remove punctuation process

Figure 5 shows the results of the remove punctuation process which was carried out at the pre-processing stage after the case folding process was carried out. The remove punctuation process is done by using a regular expression to replace all data other than characters and numbers with spaces. The results of the removal of punctuation will be followed by a tokenizing process or a process that will break the sentence into word order. By doing the tokenizing process, we can find out which words are dominant and have the highest frequency of occurrence in the document, as shown in Figure 6.
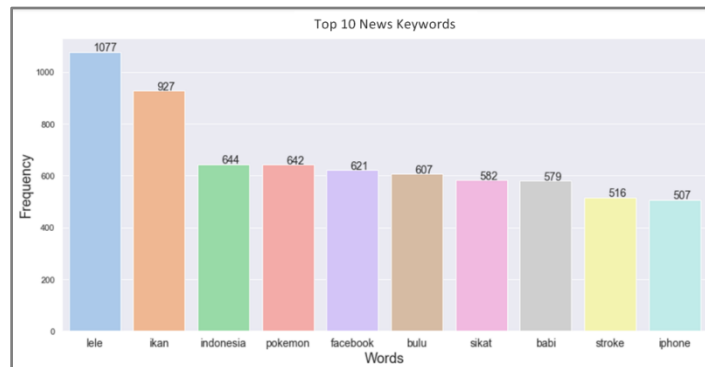
Figure 6. Top ten (10) news keywords

Figure 6, shows the result of the top ten (10) keywords that are dominant and have the highest frequency of occurrence in the document. It is what we have from doing the tokenizing process. The word "lele" have 1077 frequency in all documents, the word "ikan" have 927 frequency in all documents, the word "Indonesia" have 644 frequency in all documents, and the word "pokemon" have 642 frequency in all documents, and so on. Those keywords help recognize the main topics discussed.

After going through the tokenizing process, the next step is a stop word removal process which aims to eliminate words with a high frequency of occurrence but do not have a specific meaning that represents a class of news data. In this case, the stop word removal process is carried out using an Indonesian stop word list from Talla F.Z [12].

The last stage, after the stop word removal process, is to recombine various words that have been processed in the previous stage, so that they return to a unified whole sentence and become data that will be processed for the next processing stage.

The data that has gone through the pre-processing stage will then be extracted using the trigram method with naïve Bayes as the algorithm for the modeling process. In these steps, the author uses multinomial naïve Bayes because its model has been designed to determine term frequency i.e. the number of times a term occurs in a document. Considering the fact that a term may be pivotal in deciding the sentiment of the document, this property of this model makes it a decent choice for document classification. Also, term frequency helps decide whether the term is useful in our analysis or not. Sometimes, a term may be present in a document many times which increases its term frequency in this model but at the same time, it may also be a stopword that potentially adds no meaning to the document but possesses a high term frequency so, such words must be removed first to gain better accuracy from this algorithm [15]. Based on the results of multinomial calculations, naive Bayes and the results are known to have performance values as shown in the Figures 7.
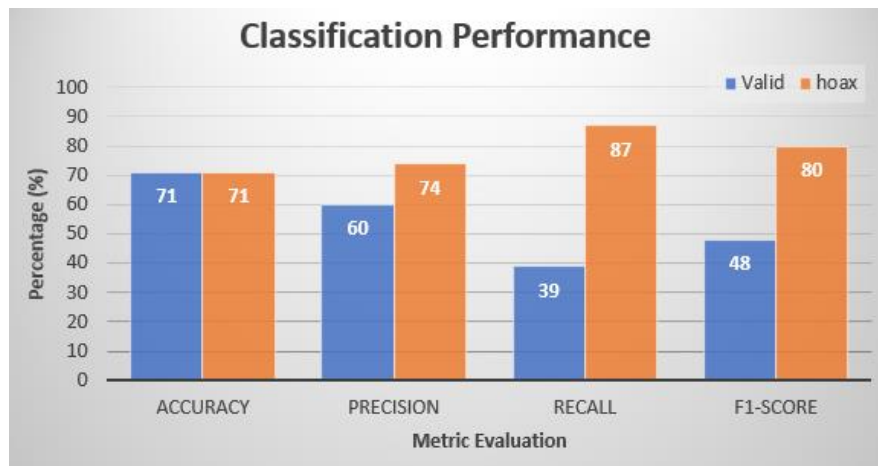
Figure 7. Metric evaluation

Figure 7 shows the performance of a machine learning model based on four (4) forms of measurement metrics, namely accuracy, precision, sensitivity, and F1-score. The performance value is obtained based on the distribution of training data and testing data with a proportion of 70:30, where 70% of the data will be used as training data used to build the model, while 30% of the other data will be used as testing data used to test the performance of the model.

Based on the results of the experiments carried out, as shown in Figure 7. the global accuracy value obtained is 71.00%. However, when viewed based on the value for each class of data, the value of precision, sensitivity (recall), and F1-Score produced shows a fairly good performance in detecting fake news (hoax) in experimental data.

The precision, sensitivity, and F1-score values obtained to determine the performance of the proposed machine learning model are calculated using the equation as described in the previous chapter. The precision value for valid news is 60%, while for hoax news it is 74%. This shows that there is a link between global accuracy results and the system in detecting fake news content (hoax) better when compared to valid news content. Likewise, the value generated on the sensitivity and F1-Score shows that the highest value is obtained in detecting the presence of fake news content (hoax) in the data. Therefore, when viewed from the value of precision, sensitivity (recall), and the resulting F1-Score, the system's prediction of fake news shows a better performance than the system's prediction of valid news.

Conclusion

Experiments related to the detection system of fake news content on Indonesian-language online media using machine learning models have been completed in this study. Naïve Bayes as the algorithm proposed as a model to predict a new invalid news content included, or fake news (hoax) produces a global accuracy value of 71%.

In addition, an experiment conducted on 600 news data using the Indonesian Hoax News Detection Dataset, obtained the respective performance values, namely a precision value of 74% for fake news detection, and 60% for valid news. A sensitivity value of 87% for hoax news is higher compared to the detection of valid news which produces a sensitivity value of 39%. For the F1-Score value, the detection of fake news (hoax) is 80%, while the detection of valid news is 48%.

Based on these results, it is known that the proposed model can be used to detect fake news content in Indonesian language online media quite well.

References

[1]   H. A. Santoso, E. H. Rachmawanto, A. Nugraha, A. A. Nugroho, D. R. I. M. Setiadi, and R. S. Basuki, "Hoax classification and sentiment analysis of Indonesian news using Naive Bayes optimization," *Telkomnika (Telecommunication Comput. Electron. Control.*, vol. 18, no. 2, pp. 799–806, 2020.

[2]   N. S. Yuslee and N. A. S. Abdullah, "Fake News Detection using Naive Bayes," in *2021*

IEEE 11th International Conference on System Engineering and Technology (ICSET), 2021, pp. 112–117.

[3]   E. Rasywir and A. Purwarianti, "Eksperimen pada Sistem Klasifikasi Berita Hoax Berbahasa Indonesia Berbasis Pembelajaran Mesin," *J. Cybermatika*, vol. 3, no. 2, pp. 1–8, 2015.

[4]   E. Zuliarso, M. T. Anwar, K. Hadiono, and I. Chasanah, "Detecting Hoaxes in Indonesian News Using TF/TDM and K Nearest Neighbor," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 835, no. 1, 2020.

[5]   E. I. Setiawan, S. Johanes, A. T. Hermawan, and Y. Yamasari, "Deteksi Validitas Berita pada Media Sosial Twitter dengan Algoritma Naive Bayes," *J. Intell. Syst. Comput.*, vol. 3, no. 2, pp. 55–60, 2021.

[6]   N. Agustina and M. Hermawati, "Implementasi Algoritma Naïve Bayes Classifier untuk Mendeteksi Berita Palsu pada Sosial Media," *Fakt. Exacta*, vol. 14, no. 4, pp. 1979–276, 2021.

[7]   A. Prasetyo, B. D. Septianto, G. F. Shidik, and A. Z. Fanani, "Evaluation of Feature Extraction TF-IDF in Indonesian Hoax News Classification," in *2019 International Seminar on Application for Technology of Information and Communication (iSemantic)*, 2019, pp. 1–6.

[8]   F. Rahutomo, I. Yanuar, and R. Andrie Asmara, "Indonesian Hoax News Detection Dataset," *Indonesian Hoax News Detection Dataset*, 2018. [Online]. Available: http://dx.doi.org/10.17632/p3hfgr5j3m.1. [Accessed: 22-Jan-2022].

[9]   I. Y. R. Pratiwi, R. A. Asmara, and F. Rahutomo, "Study of hoax news detection using naïve bayes classifier in Indonesian language," in *2017 11th International Conference on Information Communication Technology and System (ICTS)*, 2017, pp. 73–78.

[10]   H. Mustofa and A. A. Mahfudh, "Klasifikasi Berita Hoax Dengan Menggunakan Metode Naive Bayes," *Walisongo J. Inf. Technol.*, vol. 1, no. 1, p. 1, 2019.

[11]   B. Zaman, A. Justitia, K. N. Sani, and E. Purwanti, "An Indonesian Hoax News Detection System Using Reader Feedback and Naïve Bayes Algorithm," *Cybern. Inf. Technol.*, vol. 20, no. 1, pp. 82–94, 2020.

[12]   R. Yunanto, A. P. Purfini, and A. Prabuwisesa, "Survei Literatur: Deteksi Berita Palsu Menggunakan Pendekatan Deep Learning," *J. Manaj. Inform.*, vol. 11, no. 2, pp. 118–130, 2021.

[13]   A. Jain, A. Shakya, H. Khatter, and A. K. Gupta, "A smart System for Fake News Detection Using Machine Learning," in *2019 International Conference on Issues and Challenges in Intelligent Computing Techniques (ICICT)*, 2019, vol. 1, pp. 1–4.

[14]   Abdullah-All-Tanvir, E. M. Mahir, S. Akhter, and M. R. Huq, "Detecting Fake News using Machine Learning and Deep Learning Algorithms," in *2019 7th International Conference on Smart Computing Communications (ICSCC)*, 2019, pp. 1–5.

[15]   N. Smitha and R. Bharath, "Performance Comparison of Machine Learning Classifiers for Fake News Detection," in *2020 Second International Conference on Inventive Research in Computing Applications (ICIRCA)*, 2020, pp. 696–700.

[16]   M. Granik and V. Mesyura, "Fake news detection using naive Bayes classifier," in *2017 IEEE First Ukraine Conference on Electrical and Computer Engineering (UKRCON)*, 2017, pp. 900–903.

[17]   J. Singh and V. Gupta, "A systematic review of text stemming techniques," *Artif. Intell. Rev.*, vol. 48, no. 2, pp. 157–217, 2017.

[18]   Y. Pristyanto, I. Pratama, and A. F. Nugraha, "Data level approach for imbalanced class handling on educational data mining multiclass classification," in *2018 International Conference on Information and Communications Technology (ICOIACT)*, 2018, pp. 310–314.

[19]   A. F. Nugraha and L. Rahman, "Meta-algorithms for improving classification performance in the web-phishing detection process," *2019 4th Int. Conf. Inf. Technol. Inf. Syst. Electr. Eng. ICITISEE 2019*, vol. 6, pp. 271–275, 2019.

Delone and McLean Models for measuring the success of Flores University e-learning information system

Kristianus Jago Tute [1]*, Maria Adelvin Londa [2]

[1] *Information System, Technology of Informasi Faculty, Flores University*

*Corresponding Author: *jtutekristian@gmail.com*

Abstract: The University of Flores first conducted E-Learning learning amid the covid 19 pandemic, the problem faced was that most of the lecturers at the University of Flores did not understand E-learning by using information technology. The Faculty of Information Technology facilitated all faculties at the university to conduct e-training. E-Learning, in the learning process. The aim is to measure success in implementing the University of Flores E-Learning information system. The model in this study is the Delone and Mclean model to measure success with six dimensions, namely, system quality, information quality, users, service quality, user satisfaction, net benefits. The method in this research is observation, a survey using questionnaires in the form of a Likert scale, and a literature study. The data analysis technique used simple linear regression. The results showed that the Delone and Mclean models were very well used to measure the success of using e-learning. The conclusion is that the quality of information on the use percentage of 34.4%, the quality of service they use of 33.8%, the quality of the system on the satisfaction of the users is 38.6%, and the quality of the services on the satisfaction of the users 34,5%, the users on the satisfaction of the users 43,4%, use of net benefits is 58.7% and user satisfaction of net benefits is 48.8%. The higher the value of these variables, the higher the success rate of information systems using E-learning at universities will be.

Keywords: Delone and Mclean Model, information systems success, e-learning

Introduction

The Covid-19 pandemic has forced the government to issue a new policy to stop the spread of Covid-19, namely implementing public invitations to carry out physical distancing or keeping other people one meter apart and avoiding crowds and various meeting events that lead to associations [Covid-19, 2020]. The government also implements a policy to stay at home or Work From Home (WFH) and activities related to gatherings or gatherings

In this condition, all lecturers are required to replace learning using E-Learning. Various platforms are used for teaching so it needs to be supported by good learning facilities and the use of information technology [1]. Most of the lecturers at the University of Flores do not understand E-learning using information technology, so the Faculty of Information Technology at the University of Flores facilitates all faculties at the university to conduct training using E-Learning in presenting information related to the lecture process. Face-to-face online learning through applications is the most profitable thing to stop the spread of Covid-19, as well as maintain the health and safety of lecturers and students from exposure to the virus [2].

The success of the E-learning learning information system amid the Covid 19 pandemic, for Lecturers is very necessary to support all learning processes at the University of Flores. This E-learning Information System can be accessed by all lecturers at the University of Flores to include all learning materials so that the online learning process can run effectively and efficiently [3]. All students can access all information related to online learning. Measurement of information systems is needed by the University to know the added value for Higher Education [4].

The University of Flores is conducting E-Learning lessons for the first time amid the covid 19 pandemic. Researchers analyzed to measure the level of learning success using the Delone and Mclean models by measuring the success factors, namely: System Quality, Information Quality, Service Quality, Actual Use, Use Statistics, Net Benefits [5]. The role of information systems in an organization or college is no longer in doubt, its support can provide universities a competitive advantage to compete with other universities using information systems.

Methodology
Research Subject

The research subjects discussed in this study were all lecturers at the University of Flores as users of information systems. The measurement of level of success is measured based on the Delone and Mclean Model using six characteristics that aim to evaluate and find out the extent of success in implementing the University of Flores E-Learning information system.

Research Design
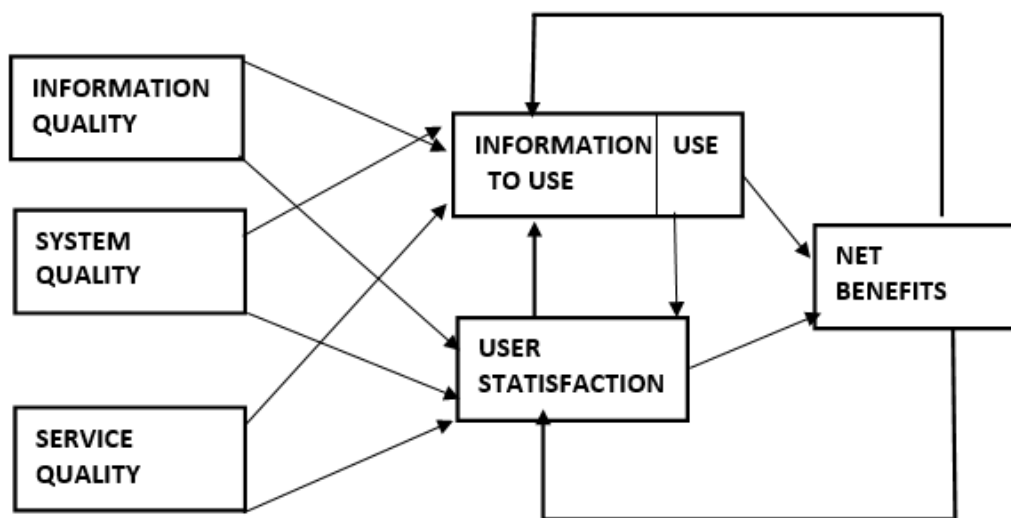
Figure 1. Information system success model

Figure 1 shows the D&M Information System Success, this model has six interconnected dimensions

Method of collecting data

Data collection methods used in this research are observation, a survey using questionnaires, literature study [6]. The analysis technique uses simple linear regression [7].

Population and sample

The population used in this study were all lecturers at the University of Flores as users of the E-Learning information system. The sampling technique is a random sampling technique in determining the number of samples using the Slovin formula so that the sample used is =204.

Validity and Reliability Test

A validity test is used to measure whether a questionnaire is valid or not from a statement that can reveal something that will be measured from the questionnaire. The research instrument is said to be valid if, the product-moment correlation coefficient exceeds 0.367, the product-moment correlation coefficient> r-table (α, ; n-2) n = number of samples, Sig value [8]. The formula for testing the validity of the instrument is as follows:

$$r = \frac{n\sum XY - (\sum X)(\sum Y)}{\sqrt{(n\sum(x)^2 - (\sum x)^2)(n\sum(y)^2 - (\sum y)^2}} \tag{1}$$

Description :
r   : Correlation coefficient r pearson
n   : Number of samples/observations
x   : Independent variable/first variable
y   : Dependent variable/second variableơ

The validity data of calculating are as follows:

Table 1. Validity test

Case Processing Summary		N	%
Cases	Valid	204	100.0
	Excluded[a]	0	0.0
	Total	204	100.0
a. Listwise deletion based on all variables in the procedure.			

Reliability Test

To calculate the reliability measurement using Cronbach's Alpha, with this technique if the reliability coefficient with r value > 0.600. For the calculation process using SPSS[9]. Cronbach's Alpha formula is as follows:

$$r11 = \left(\frac{k}{(k-1)}\right) \; (1 - \frac{\sum \sigma t^2}{\sigma t^2}) \tag{2}$$

Description :
r 11   : Instrument reliability
k   : number of questionnaire items
$\sum \sigma t^2$: number of item variants
σt2   : Total Variants

Croanbach's alha table is as follows:

Table 2. Reliability Test

Reliability Statistics		
Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
0.056	0.963	35

After compiling the questionnaire and testing the validity, reliability, and distributing questionnaires to all lecturers at the University of Flores using E-Learning, the next step is to conduct an analysis.

Results and Discussions

This study was adapted from the information system success model developed by Delone and Mclean. This study uses six variables consisting of system quality, information quality, service quality, usage, user satisfaction, and net benefits. The purpose of this study was to determine

the effect of these variables in the use of the University of Flores' E-Learning information system. This study has nine hypotheses to see which variables affect the success of the University of Flores E-Learning information system.

1.   System Quality to Use

From the results of a simple linear regression test on the system quality variable (System Quality) to the user (Use), the value is obtained from F count > T table.

Table 3. Coefficient of system quality to users

Coefficients					
Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
1 (Constant)	0.300	0.057		5.223	0.000
NZXI	0.593	0.058	0.587	10.298	0.000
a. Dependent Variable: NZY1					

From the ANOVA test by paying attention to R Square, the regression coefficients that are generated with t-count values of 5.223 and 10.298 with a significance of 0.000, show that the quality of the system affects the use of each with a significant coefficient.

Table 4. ANOVA test

ANOVAa					
Model	Sum of Squares	df	Mean Square	F	Sig.
1 Regression	30.958	1	30.958	106.059	0.000b
Residual	58.964	202	0.292		
Total	89.922	203			
a. Dependent Variable: ZUSE					
b. Predictors: (Constant), ZKS					

The results of the ANOVA test show F count 106,059 with a significant 0.000, then compared with F table with df numerator of 1 and df denominator 202 with a significant level of 0.05 is 3.04, it is very clear that F count > F table 4. So it can be concluded that the quality of the system is very influential on its use. This is in line with research conducted by [10] if the quality of the system is good, then user satisfaction will increase.

2.   Information Quality on Use

Table 5. System quality to users

Coefficientsa					
Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
1 (Constant)	0.351	0.057		6.108	0.000
NZX2	0.538	0.058	0.549	9.334	0.000
a. Dependent Variable: NZY1					

By looking at the R Square value with a model of 0.301 or 30.1%, the quality of information does not significantly affect the use. The resulting model with t-count values of 6.108 and 9.334 with a significance of 0.000 shows that each coefficient has no significant effect on the use. The quality of information with a B value of 0.351 while the effect on the use of 0.538.

Table 6. ANOVA test

Model		Sum of Squares	df	Mean Square	F	Sig.
ANOVAᵃ
1	Regression	27.098	1	27.098	87.130	0.000b
	Residual	62.824	202	0.311		
	Total	89.922	203			

a. Dependent Variable: ZYU

b. Predictors: (Constant), ZX2KI

Table 7. Summary model

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	0.549ᵃ	0.301	0.298	0.55768	0.301	87.130	1	202	0.000

a. Predictors: (Constant), ZX2KI

b. Dependent Variable: ZYU

The results of the ANOVA test F Count = 87,130 with a significant 0.000, then compared with Ftable with df the numerator of 202 and df of the denominator 203 at 0.05, the value of which is 3.04 so it can be concluded that the information quality variable has an effect on usage but is less significant on usage because by looking at the R-value Square of 0.301. This result shows that the higher the quality of the information used, the higher the use of e-learning amid the COVID-19 pandemic [11].

3. Quality of service to use

Table 8. ANOVA test

Model		Sum of Squares	df	Mean Square	F	Sig.
ANOVAᵃ
1	Regression	30.412	1	30.412	103.230	0.000b
	Residual	59.510	202	0.295		
	Total	89.922	203			

a. Dependent Variable: ZKUse

b. Predictors: (Constant), ZKPEL

From the results of the F test, the calculated F value is 103,230 with a significant level of 0.000, then compared with the F table with df numerator 1 and df free denominator 202 with a significant level of 0.05 whose value is 3.04. it seems very clear that F count = 103.230 > from F table 3.04, so it can be concluded that the quality of service is very influential on usage.

Table 9.Coefficients

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	0.322	0.056		5.717	0.000
	ZKPEL	0.574	0.056	0.582	10.160	0.000
a. Dependent Variable: ZKUse						

The resulting model with a t value of 5.717 > 0.056 and a significance of 0.05 (0.000 < 0.05) so it can be concluded that the quality of service is very influential on usage. The results of this study are in line with research conducted by [11] the better the quality of service in the learning process with e-learning, the better an increase in the use of e-learning.

4.   System Quality on user satisfaction

Table 10. Summary Model

Model Summary^b

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	0.621^a	0.386	0.383	0.50521	0.386	126.817	1	202	0.000
a. Predictors: (Constant), ZKS									
b. Dependent Variable: ZKP									

By looking at the correlation R of 0.621 the effect of system quality on user satisfaction. The coefficient of determination is 0.383 with the understanding that the effect of the independent variable on the dependent variable is 38.3%, while the rest is influenced by other variables.

Table 11. ANOVA test

ANOVA^a

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	32.369	1	32.369	126.817	0.000^b
	Residual	51.558	202	0.255		
	Total	83.927	203			
a. Dependent Variable: ZKP						
b. Predictors: (Constant), ZKS						

Table 11 shows that the calculated F value = 126,817 > from Ftable 3.04 with a probability level of 0.000 < 0.05, regression can be used to predict the participation variable.

Table 12.Coefficients

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	0.309	0.054		5.762	0.000
	ZKS	0.606	0.054	0.621	11.261	0.000
a. Dependent Variable: ZKP						

Table 12 shows the value of t arithmetic = 5.762 with a significant value of 0.000 <0.05, then rejecting H0 and accepting H1 means that there is an influence between system quality and user quality. The results of research conducted by[12] show that the better the quality of the e-learning system used, the better the satisfaction of e-learning users at the University of Flores.

5.   Information Quality on User Satisfaction

Table 13. Summary Model

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	0.404[a]	0.163	0.159	0.58970	0.163	39.345	1	202	0.000
a. Predictors: (Constant), ZKI									
b. Dependent Variable: ZKP									

The correlation value R = 0.404 shows the percentage of the influence of information quality on user satisfaction by looking at the coefficient of determination R Square of 0.163, implying that the influence of information quality on user satisfaction is 16.3%, not too significant.

Table 14. ANOVA test

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	13.682	1	13.682	39.345	0.000[b]
	Residual	70.245	202	0.348		
	Total	83.927	203			
a. Dependent Variable: ZKP						
b. Predictors: (Constant), ZKI						

From the ANOVA test and paying attention to R Square 0.163, which is less significant, the resulting coefficient is not good, so it cannot be used for analysis. So it can be concluded that the quality of information does not significantly affect the user. This is because the University of Flores is already using e-learning, but only some of the lecturers use e-learning. This research is in line with [13].

6.   Service Quality on User Satisfaction

Table 15. Summary Model

Model Summary^b

Mo del	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	0.58 8^a	0.345	0.342	0.52155	0.345	106.5 34	1	202	0.000

a. Predictors: (Constant), ZKP

b. Dependent Variable: ZKPM

The correlation value R = 0.588 shows the percentage of the influence of service quality on user satisfaction with a coefficient of determination R Square = 0.345, that the effect of service quality on user satisfaction is 34.5% and the rest is influenced by other variables.

Table 16. ANOVA test

ANOVA^a

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	28.979	1	28.979	106.534	0.000^b
	Residual	54.948	202	0.272		
	Total	83.927	203			

a. Dependent Variable: ZKPM

b. Predictors: (Constant), ZKP

The calculated F value = 106.534 with a probability of 0.000 <0.05, then the regression model can be used. This value is then compared with the value of Ftable 3.04. It seems very clear that Fcount = 106,534 is greater than Ftable = 3.04, so it can be concluded that service quality is very influential on user satisfaction. This research is in line [14], showing that there is an influence between service quality on user satisfaction.

7.   Users to User Satisfaction

Table 17. Summary Model

Model Summary^b

Mo del	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	0.658 ^a	0.434	0.431	0.48511	0.434	154.6 38	1	202	0.000

a. Predictors: (Constant), NZP

b. Dependent Variable: NZKP

The correlation value R = 0.658 shows the percentage of the effect of the use on user satisfaction with a coefficient of determination R Square = 0.434, which means that there is a 43.4% effect of usage on very good user satisfaction and the remaining 56.6% is influenced by other variables.

Table 18. ANOVA test

ANOVAa						
Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	36.391	1	36.391	154.638	0.000b
	Residual	47.536	202	0.235		
	Total	83.927	203			
a. Dependent Variable: NZKP						
b. Predictors: (Constant), NZP						

The calculated F value = 154.638 with a probability level of 0.000 <0.05, then this model can be used to predict the participation variable. The value of t count = 5.690 with a significance of 0.000 <0.05, rejecting H0 and accepting H1. It can be concluded that the user is very influential on user satisfaction. The results of this study are in line with research conducted by [15] there is a user influence on user satisfaction.

8. Use against Net Benefits

Table 19. Summary Model

Model Summaryb									
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	0.766a	0.587	0.585	0.39566	0.587	286.732	1	202	0.000
a. Predictors: (Constant), NZP									
b. Dependent Variable: NZNB									

The value of R = 0.766 indicates that the effect of the use on net benefits is quite significant with a coefficient of determination R Square = 0.587, that the effect of the use on net benefits of 58.7% is very good to use and the remaining 41.3% is influenced by other variables

Table 20. ANOVA test

ANOVAa						
Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	44.888	1	44.888	286.732	0.000b
	Residual	31.623	202	0.157		
	Total	76.511	203			
a. Dependent Variable: NZNB						
b. Predictors: (Constant), NZP						

From the Anona test by looking at the value of R square = 0.587, then the calculated F value = 286.732 with a probability of 0.000 <0.05, the regression model can be used. This value is then compared with the value of Ftable 3.04. It seems very clear that Fcount = 286.732 is greater than Ftable = 3.04, so it can be concluded that the use has a great effect on net benefits. The results of this study conducted by [16], there is an effect of the use on net benefits.

9.   User Satisfaction with Net Benefits

Table 21. Summary Model

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	0.699[a]	0.488	0.486	0.44020	0.488	192.839	1	202	0.000

a. Predictors: (Constant), NZK

b. Dependent Variable: NNB

The value of R = 0.699 indicates that the percentage of the effect of user satisfaction on net benefits with a coefficient of determination R Square = 0.488 or 48.8% of user satisfaction on net benefits and the remaining 51.2% is influenced by other variables.

Table 22. Anova test

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	37.368	1	37.368	192.839	0.000[b]
	Residual	39.143	202	0.194		
	Total	76.511	203			

a. Dependent Variable: NNB

b. Predictors: (Constant), NZK

The value of R = 0.699 indicates that the percentage of the effect of user satisfaction on net benefits with a coefficient of determination R Square = 0.488 or 48.8% of user satisfaction on net benefits and the remaining 51.2% is influenced by other variables. The calculated F value = 192.839 with a probability level of 0.000 <0.05, then this model can be used to predict other variables. Thus the value of F count = 192.839 > from F table = 3.04, then user satisfaction is very influential on net benefits. The results of this study are in line with research conducted by [17] showing that user satisfaction has a significant effect on net benefits.

Conclusion

This study uses the Delone and Mclean model to measure the success of the University of Flores E-Learning information system by looking at the influence of each variable that exists in the Delone and Mclean model, the results of the analysis of the success of its application based on the discussion above can be concluded that: variables that have a large and positive effect on the success of the E-Learning information system at the University of Flores, namely the variable quality of information on use with a percentage of 34.4%, service quality on use 33.8%, system quality on user satisfaction 38.6%, service quality to user satisfaction 34.5%, user satisfaction to user satisfaction 43.4%, usage to net benefits 58.7% and user satisfaction to net benefits 48.8%. The higher the value of these variables, the higher the success rate of the information system using E-learning at the University of Flores. The Delone and Mclean models are very successful in measuring the information system and it is hoped that all Flores University lecturers are required to use e-learning.

References

[1]   M. Meidiawani, M. Misnaniarti, and R. A. Syakurah, "Kepuasan Pengguna Aplikasi E-Ppgbm Berdasarkan Model Kesuksesan Delone -Mclean," *Prepotif J. Kesehat. Masy.*, vol. 5, no. 1, pp. 96–102, 2021, doi: 10.31004/prepotif.v5i1.1323.

[2]   Mustakim, "Efektivitas Pembelajaran Daring Menggunakan Media Online Selama Pandemi Covid-19 Pada Mata Pelajaran Matematika the Effectiveness of E-Learning Using Online Media During the Covid-19 Pandemic in Mathematics," *Al asma J. Islam. Educ.*, 2020.

[3]   K. Sara, F. L. Witi, and A. Mude, "Implementasi E-Learning Berbasis Moodle di Masa Pandemi Covid 19," *J. Adm. Educ. Manag.*, vol. 3, no. 2, pp. 181–189, 2020, doi: 10.31539/alignment.v3i2.1813.

[4]   L. Meilani, A. I. Suroso, and L. N. Yuliati, "Evaluasi Keberhasilan Sistem Informasi Akademik dengan Pendekatan Model DeLone dan McLean," *J. Sist. Inf. Bisnis*, 2020.

[5]   S. Rahayu, "Pengaruh Kesuksesan Sistem Dapodikdasmen Kota Palembang Menggunakan Model Delone dan McLean," *J. Sains, Teknol. dan Ind.*, 2020, doi: 10.24014/sitekin.v16i2.8270.

[6]   Sugiyono, *Metode Penelitian Kuantitatif, Kualitatif dan R&D*, Cetakan ke 21. Bandung: Alfabeta, 2015.

[7]   Badan Pusat Statistik Indonesia, "Statistik Indonesia Tahun 2020," *Stat. Indones. 2020*, 2020.

[8]   A. Arsi, "Realibilitas Instrumen Dengan Menggunakan Spss," *Validitas Realibilitas Instrumen Dengan Menggunakan Spss*, 2021.

[9]   S. Zein, L. Yasyifa, R. Ghozi, E. Harahap, F. Badruzzaman, and D. Darmawan, "Pengolahan dan Analisis Data Kuantitatif Menggunakan Aplikasi SPSS," *J. Teknol. Pendidik. dan Pembelajaran*, 2019.

[10]   K. D. P. Novianti, "Analisis Evaluasi E-learning Menggunakan Integrasi Model D&M dan UTAUT," *Techno.Com*, 2019, doi: 10.33633/tc.v18i2.2217.

[11]   D. Suhendro, "Pengaruh Kualitas Sistem, Kualitas Informasi, Kualitas Pelayanan Dan Ekspektasi Kinerja Terhadap Kepuasan Pengguna Dalam Penerapan Sistem Teknologi Informasi Pada Koperasi Di Kota Pematangsiantar," *Jurasik (Jurnal Ris. Sist. Inf. dan Tek. Inform.*, vol. 1, no. 1, p. 33, 2017, doi: 10.30645/jurasik.v1i1.6.

[12]   G. P. L. Permana and N. W. Mudiyanti, "Analisis Faktor Kesuksesan Implementasi Aplikasi Sistem Keuangan Desa (Siskeudes) Dengan Menggunakan Model Kesuksesan Sistem Teknologi Informasi Diperbarui Oleh Delone Dan Mclean Di Kabupaten Gianyar," *Krisna Kumpul. Ris. Akunt.*, vol. 13, no. 1, pp. 75–85, 2021, doi: 10.22225/kr.13.1.2021.75-85.

[13]   A. P. Utomo *et al.*, "Model Evaluasi Integrasi Kesuksesan dan Penerimaan Sistem Informasi E-learning Universitas," *Ikraith Inform.*, 2018.

[14]   A. S. B. Wahyudi and W. Wardiyono, "Evaluasi Pemanfaatan Sistem Informasi Kasus Dengan Model Information System Success Delone & Mclean Di Lembaga Bantuan Hukum Jakarta," *Bibl. J. Ilmu Perpust. dan Inf.*, 2018, doi: 10.33476/bibliotech.v3i2.914.

[15]   D. Sarumaha and J. J. C. Tambotoh, "Evaluasi Penerimaan Fintech Berdasarkan Model DeLone-McLean dan Penerimaan Teknologi," *J. Nas. Teknol. dan Sist. Inf.*, 2022, doi: 10.25077/teknosi.v7i3.2021.199-210.

[16]   A. M. Nuur, N. Wanti Wulan Sari, and Faldi, "Analisis Kesuksesan E-Learning Google Classroom pada Universitas Mulia Kampus Samarinda Dengan Metode Delone and Mclean," *Metik*, 2019.

[17]   J. M. Hudin, Y. Farlina, and D. Pribadi, "Pengukuran Tingkat Kesuksesan Sistem Informasi Ruang Mahasiswa Pada Universitas BSI PSDKU Sukabumi Dengan Model Delone Dan Mclean," *Indones. J. Bus. Intell.*, 2019, doi: 10.21927/ijubi.v2i2.1106.

Business transformation from connectivity to digital: case study "PT. MNO Indonesia"

Itsnanta Muhammad Fauzan [1], *Amirul Ihsan* [2], *Farhan Annur Mahmudi* [3], *Amirsyah Rayhan Mubarak* [4], Catur Apriono [5]*

[1,2,3,4,5] *Universitas Indonesia, Indonesia*

*Corresponding Author: *catur@eng.ui.ac.id*

Abstract: The telecommunications industry plays a role in meeting the needs of digital connectivity services to the community, especially in line with the community's need for internet connections which is increasing rapidly. As one of the telecommunications service operators in Indonesia, "PT. MNO Indonesia" must meet this need to get maximum profit and the best service to the customers with the right business strategy. However, the current connectivity business model is still the mainstay, with revenue levels that tend to stagnate and even experience a decline. The provision of network services and investment in infrastructure development are still the main concerns, so a change in business strategy is needed. This study proposes several recommendations for business strategies to increase revenue based on reviewing and analyzing financial data and statistics for Indonesian and international telecommunications businesses. In this study, financial report data and digital sector statistical data were analyzed and divided into four aspects: traffic, finance, network development, and digital service with the min-max method, standard deviation, and average. From the analysis conducted in this research, "PT. MNO Indonesia" needs to transform its business into digital business model.

Keywords: business transformation, connectivity business model, digital

How to Cite: I. M. Fauzan, A. Ihsan, F. A. Mahmudi, A. R. Mubarak and C. Apriono, "Business transformation from connectivity to digital: case study "PT. MNO Indonesia," Matrix: Jurnal Manajemen Teknologi dan Informatika, vol. 12, no. 2, pp. 79-90, 2022.

Introduction

A digital business is a business that uses technology as an advantage in internal and external operations. Another definition of digital business is the most advanced business today, which is expected to continue to grow every year and increase sharply [1]. Digital technology has revolutionized the entire joint of life, giving birth to a new civilization. All service companies are competing to provide digital version services for consumers and are committed to maximizing the use of digital technology by presenting a variety of service digitization solutions that are increasingly integrated.

The development of information technology and telecommunications greatly impacts the business sector. Business systems that previously used traditional patterns were transformed into digital by following the needs of people who demand speed and efficiency in service and buying and selling goods. Digital businesses also provide benefits in ease of transactions by connecting sellers and buyers using online media. It makes many companies that create platforms (websites or startups) to support the business.

5G technology is a 5th generation cellular network that is the wireless standard after 1G, 2G, 3G, and 4G networks [2]. 5G technology is designed to connect all objects, machines, and devices simultaneously by providing high data speed capabilities up to 20 times faster than the previous generation (4G), i.e., based on IMT-2020 reaching 20 Gbps. 5G technology is also designed for services with large network capacity and wide bandwidth as in the mmWave frequency band spectrum to provide faster responses with very low latency so that users as a whole get a more uniform experience despite the user's moving conditions.

5G technology was officially launched in Indonesia in early 2021 through a press conference between the Ministry of Information and Communication (Kominfo) and one of the operators. The operation of 5G networks is a form of accelerating digital transformation in Indonesia. 5 G ser-vices connect communication between people and provide human integration services with ma-chines. For 5G services to run optimally in Indonesia, it requires the allocation of the frequency spectrum in three layers, namely low band, middle band, and high band. Each layer can later be used according to the services needed. Currently, operators are preferred to use the middle band layer for mobile broadband development. Together with industry players, the government pre-pares ecosystems and devices such as smart cities, e-commerce, and the Internet of Things (IoT) in the industry.

The development of telecommunication technology to the fifth generation (5G) encourages mobile network operator companies to adapt by carrying out digital transformation. The Covid-19 pandemic is one factor driving the acceleration of digital transformation, where people switch to utilizing telecommunication technology to carry out various activities. This condition makes the acceleration of cell operator companies follow the digital transformation process. Mobile operator companies in the last ten years (decade) are still relying on connectivity businesses model with stagnant revenue levels and even tend to fall [3].

Continuous advances in high-speed mobile services and internet connectivity between devices continue to drive competition and innovation in the telecommunications service provider sec-tor. Much of the main focus of the business rests on providing faster data services, including high-resolution video transmission services. Other driving factors of the telecommunications business are the demand for increased connectivity and multiplication. This condition places the mo-bile op-erator company as a network provider (dump-pipe), where other parties such as application pro-viders, OTT, and others manage the implementation of network services.

From an investment point of view, 5G technology infrastructure requires very large costs, including coverage of the 5G area of narrower telecommunication devices, contributing to development costs. In addition, the deployment of fiber optics as a backbone network also increases the investment costs of mobile operators. Network-related capital expenditures are expected to increase by 60% from 2020 to 2025, resulting in a two-fold increase in total cost ownership (TCO) [4]. Nonetheless, mobile operators must upgrade existing 4G networks to 5G networks to cope with the growing demand for data year after year [4]. In addition, the service innovations, e.g., machine-to-machine communication and virtual reality that cannot be supported by 4G technology, encourage operators to build 5G networks [5]. As one of the mobile operator companies in Indo-nesia, "PT. MNO Indonesia" has a rather high investment burden for the development (CAPEX) of telecommunication networks while revenue conditions have not increased significantly. During the last five years, from 2016 to 2021, the CAPEX average of 15% was allocated for mobile network improvements. Revenue conditions were no longer rising in double digits as in the previous year. This condition provides a tendency that is not good for running the company's business, so the right strategy is needed to maintain the sustainability of both the business and the provision of services for customers.

Ideally, communication needs in Indonesia should be able to reach all regions from Sabang to Merauke, providing equitable access to data communication. However, the penetration rate of telecommunication infrastructure is still uneven. About 40% of rural areas, especially in the Forefront, Remote, and Disadvantaged areas, still do not get the reach to telecommunication services [6]. The Government of Indonesia supports digital development in Indonesia through The Minister of Communication and Information Regulation No. 2 of 2021, concerning the acceleration of digital transformation [7] with the condition of the Eastern Indonesia Region (KIT), which has 11 prov-inces with eight industrial estate locations as targets for the development of mobile tele-communi-cation networks. One of the programs launched in the '3T' area ("Terdepan, Terluar & Terting-gal") was inducted in 2015-2020. From 2021 to 2022, the government continues to provide BTS (Base Transceiver Station) in 7,904 Villages, where the largest allocation is in the Eastern Indone-sia Region [8].

In some of the scientific literature that has been reviewed, the discussion in this study is about business transformation to increase revenue potential and maintain sustainability. Based on literature related to research topics reviewed from the point of view of background and research methods, discussions are still limited to the development of technology (smart city, IoT,

and wire-less sensors), investment, 5G infrastructure, and industry changes to digital, where no specific discussion with business strategy is associated [9-28].

This research focuses on preparing a proposed recommendation for "PT. MNO Indonesia" to transform the telecommunications business from Legacy /Connectivity towards Digital Business to be able to survive to become the largest operator in Indonesia with a decent level of profitability, accelerate the development of targeted 5G infrastructure, make investments and innovations on the platform. Digital and automation to create new business opportunities.

Methodology

To formulate recommendations for the business strategy, the collected data was processed and analyzed using quantitative and qualitative secondary data approaches. These data were divided into four aspects: traffic, financial, network deployment, and digital services as summarized in Table 1. Figure 1 shows a flow chart of the research carried out in this study.
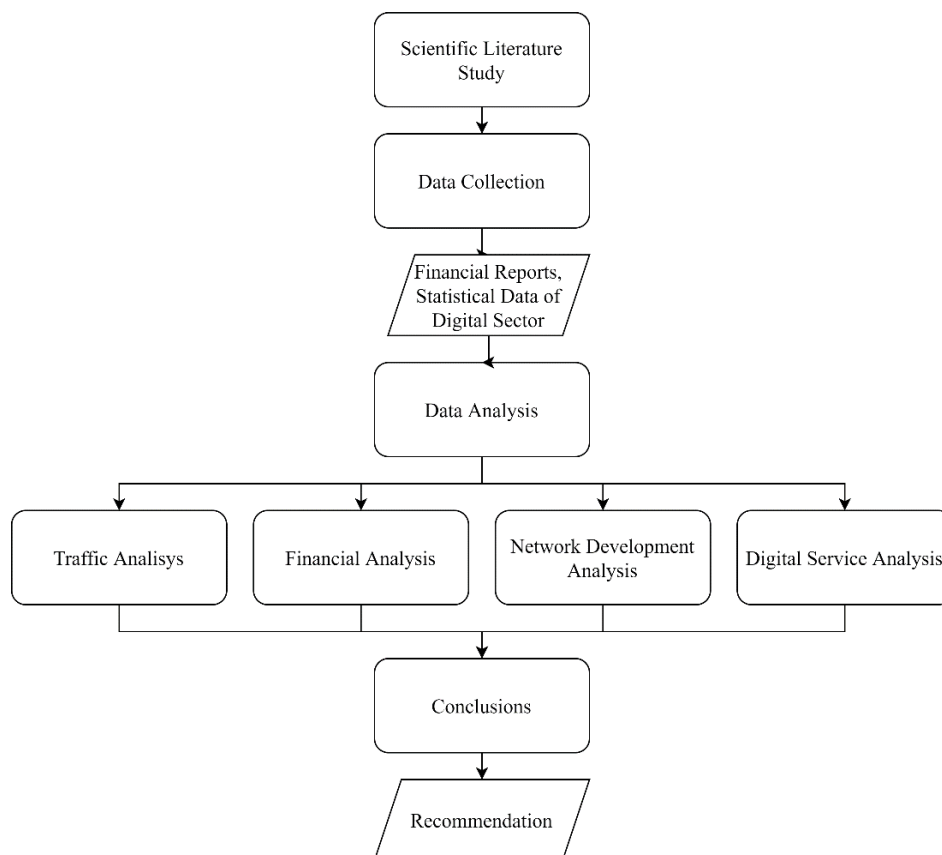

Figure 1. Research flowchart

Table 1. Data categories to be analyzed

Aspect	Parameter	Description
Traffic	Connections	Data of customer total measured connections to the digital services of "PT. MNO Indonesia".
	ARPU	Data of average income of the digital services of "PT. MNO Indonesia" per connected customer.
Financial	Data Revenue	Data of total annual income of "PT. MNO Indonesia".
	Data CAPEX/Revenue Ratio	Data of percentage ratio of "PT. MNO Indonesia" income with annual CAPEX.
	Data Net Profit	Data of total annual net profit of "PT. MNO Indonesia".
Network Deployment	Internet Access Distribution	Data on the level of connectivity of service users to internet access per region in Indonesia at "PT. MNO Indonesia".
Digital Services	ICT Market	Market data from the ICT business in Indonesia in recent years and forecasts for the next few years.
	Market Shifting	Data pointer shifting market model in Indonesia
	Digital Platform Investation Data	Digital service investment data on digital platforms by "PT. MNO Indonesia".

1. Scientific Literature Study

At this stage, we review scientific articles and popular news, which are used as material to review the research topics to obtain problems and urgency in the telecommunications industry.

2. Data Collection

This research considers secondary data, both quantitatively and qualitatively, from various sources related to the telecommunications business. The data used are the financial report data of "PT. MNO Indonesia", statistical data on the digital sector of Indonesia, as well as comparative data from other countries. These data will be analyzed and used to support the recommendations at the end of the study. The data used is limited to the range of the last ten years to retain relevancy.

3. Data Analysis

The data analysis process is the next step after data collection. The four categories of existing data can be reviewed on statistical parameters such as the average parameter, standard deviation, and min-max. The analysis was done by comparing the data values between periods in one particular aspect, as shown in Table 2. Based on the analysis obtained, it will be used to provide conclusions in determining whether or not a business transformation is necessary.

4. Conclusions

This research aims to produce business strategy recommendations from the analyzed data. The results of the analysis of this study are in the form of recommendations for "PT. MNO Indonesia" and other companies both nationally and internationally with appropriate conditions.
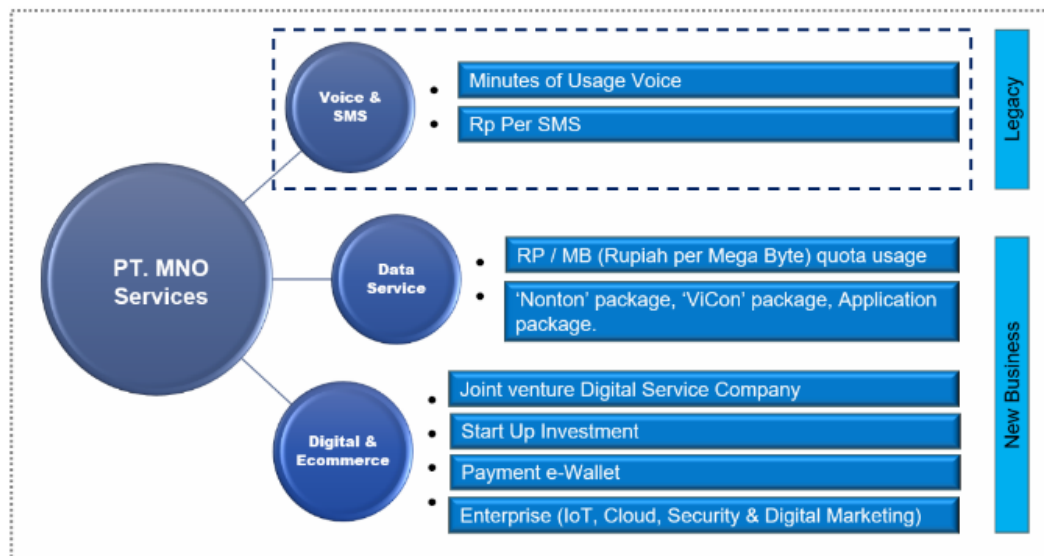
Figure 2. Research mind map

Figure 2 shows a mind map illustrating the "PT. MNO Indonesia " services with the legacy (legacy) and the new business potential in the digital transformation process. The beginning of the telecommunications business was voice and SMS services for mobile communications, and the development of cellular technology provided additional services, namely data services. Previously, data services only served the needs of the internet, video calls, and chats. Still, it has been devel-oping since the existence of other services from sharing applications that are very easy to find in the smartphone system. This condition makes the business in "PT MNO Indonesia " remain com-petitive and sustainable by developing previous services in 4G technology (digital payment, video conference platforms, watching packages, and application packages), new businesses in 5G tech-nology (IoT, security, cloud) and digital marketing as well as investing with other digital compa-nies.

Results and Discussions

There are 4 (four) aspects of elaboration needed to pay attention to the transformation points of "PT. MNO Indonesia", namely aspects of traffic, finance, network deployment, and dig-ital services.

Figure 3 shows a trend graph of Total Cellular Connection and total ARPU from "PT. MNO Indonesia " from 2012 to 2021. Total Cellular Connection tends to stagnate at 178 million connec-tions in Q4-2021 and does not increase compared to the previous four years (181 million connec-tions in Q2-2017). It can be seen that the increasing number of users will support the probability of traffic growth, as well as indirectly also increase revenue for the operator company, so "PT. MNO Indonesia" needs to maintain or increase provider users with products or services that are affordable, attractive, and up-to-date. The potential that digital business will continue to grow, with services from 5G technology services will bring interest and positive impact for users and providers.

Figure 3. Total cellular connection "PT. MNO Indonesia" [3]

Figure 4 shows the average revenue per user (ARPU) of the company "PT. MNO Indonesia ". It indicates that the total ARPU has not increased since six years ago, which was 3.01 USD in Q4-2015 to 3.01 USD in Q4-2021. This condition encourages the company to grow user connections by growing the potential for new service penetration. With the increase in the ARPU value, the revenue per user will provide value growth for the company.
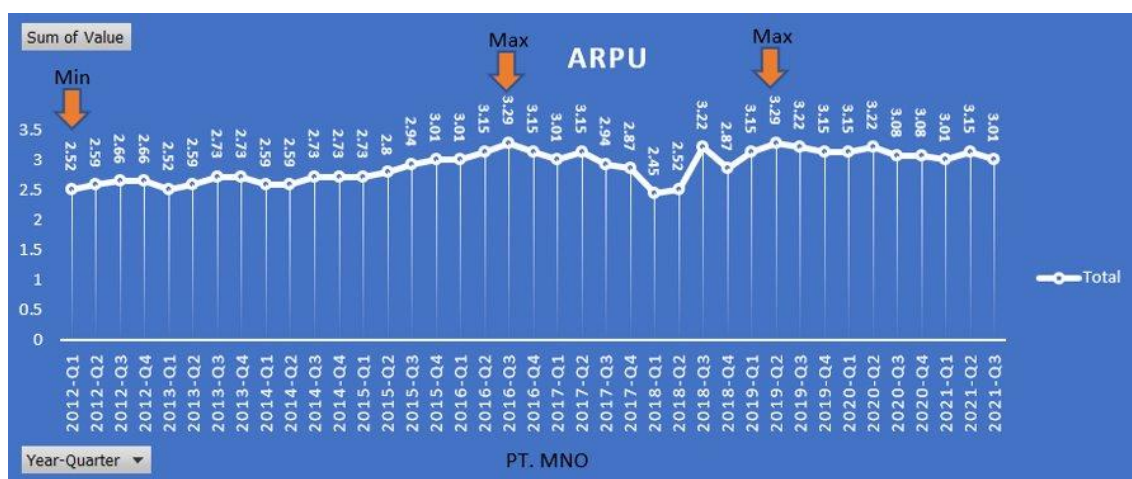

Figure 4. ARPU by connection "PT. MNO Indonesia" [3]

Figure 5 shows revenue trends for the period 2012 to 2021. It has a positive business trend, and it is observed from the results that revenue has continued to grow over the last three years, namely from 2.1 billion USD in Q2 -2018 to 2.5 billion USD in Q3-2021. The company's total revenue continues to grow, although it should be noted that its revenue has been saturated in the last two years. Until now, the company is still the largest provider in Indonesia. Similar to the traf-fic aspect, the financial aspect will grow by increasing the penetration of the cellular telecommuni-cations market from the company "PT. MNO Indonesia "and maintaining stability in terms of de-velopment and operational costs of the company.

Figure 5. Revenue "PT. MNO Indonesia" [3]

Figure 6 shows a graph of the percentage ratio of the CAPEX/revenue value. The CAPEX/revenue ratio value shows relatively stable results (between the last two years, namely 15.33% in Q2-2019 and 15.81% in Q3-2021). With the shifting streamline from legacy business to digital, any development to maintenance action will be a big factor that has the potential to reduce revenue. It shows that the company needs to continue to maintain it, extend assets and deploy-ment from the access infrastructure sector to the core in a stable manner, and adjust to the in-come earned by the company.
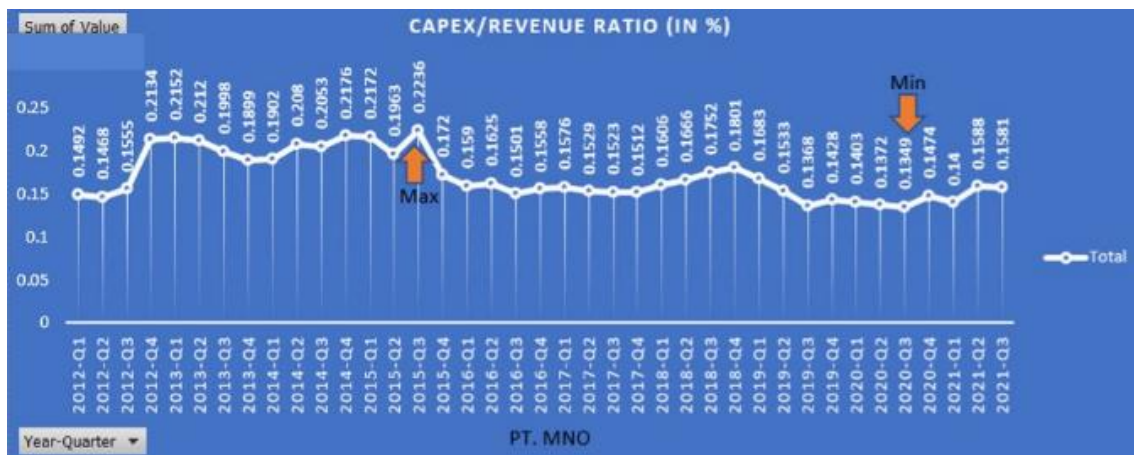
Figure 6. CAPEX to revenue ratio "PT. MNO Indonesia" [3]

Figure 7 shows a graph of the net profit value. The data found that the company's net profit was in a stable range of values, reaching 507 billion USD in Q4-2018. After that, it was at 470 billion USD, and the latest data in Q3-2021 Net Profit was 467 billion USD. The company "PT. MNO Indonesia" earns relatively healthy profits due to slowly growing profits, so the company's market in the telecommunications business is running well. These three parameters show that the company's business is still healthy. The increase in revenue and net profit needs to be main-tained so that the company's condition continues to run.

Figure 7. Net profit "PT. MNO Indonesia" [3]

This aspect shows how the penetration of internet access in Indonesia is. Indonesia still has gaps in telecommunications network infrastructure, especially for provinces with limited last-mile physical networks. Figure 8 shows internet access in the provinces in Indonesia. Based on BPS data, most penetration of Internet Access distribution is dominated by big cities in Indonesia. These cities including DKI Jakarta, D.I. Yogyakarta, Riau, East Kalimantan, and Bali already have the infrastructure and a high population. Meanwhile, several provinces in the Eastern Indonesia Region (KIT) are included in the distribution of Internet access from the lowest, namely Papua, NTT, North Maluku, West Sulawesi, and Maluku. The potential shown should serve as a bench-mark so that the company can expand its network development to become larger in cities that have the potential to increase revenue for the company, especially for 5G services.
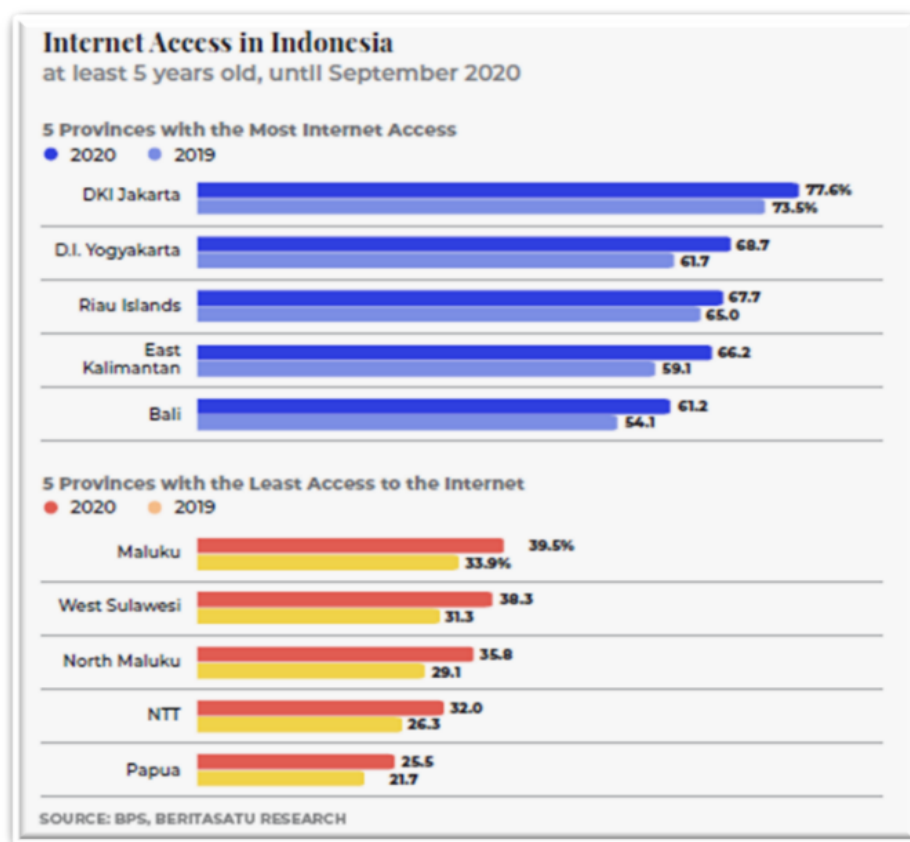

Figure 8. Distribution of internet access in Indonesia to Provinces [6]

This aspect illustrates that digital business is expected to trigger new penetration for companies. The current trend in almost all countries has the same portrait related to the telecommunications business. The presence of digital services provides productive potential in the future and the saturation of the pure telecommunications business. Figure 9 shows a graph of the ICT market in Indonesia, where the telecommunications sector tends to be saturated, and the cyber security sector has not grown significantly. Still, the trend will become an important factor in the future, and the Digital Service sector is a factor that is expected to grow rapidly in the next few years.
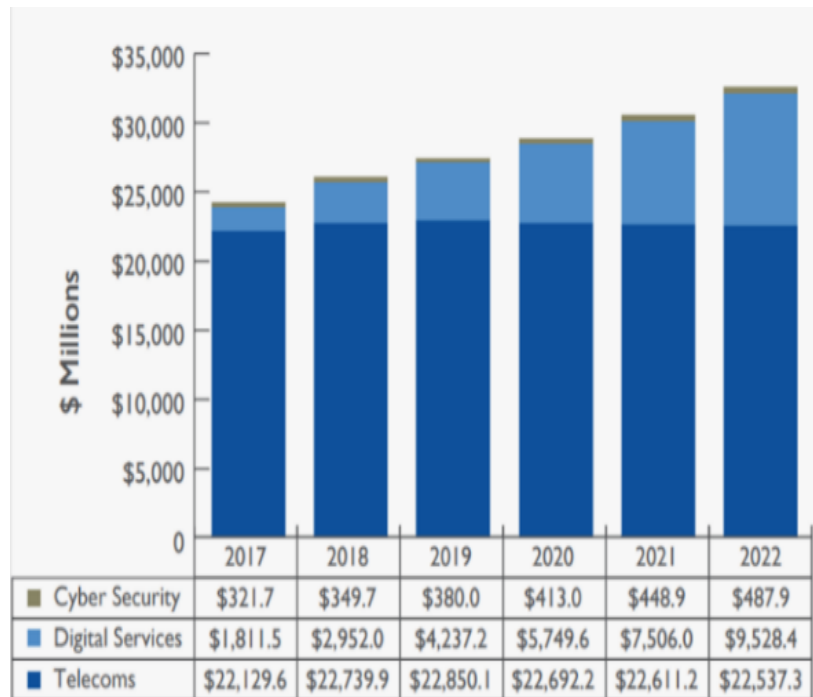

	2017	2018	2019	2020	2021	2022
■ Cyber Security	$321.7	$349.7	$380.0	$413.0	$448.9	$487.9
■ Digital Services	$1,811.5	$2,952.0	$4,237.2	$5,749.6	$7,506.0	$9,528.4
■ Telecoms	$22,129.6	$22,739.9	$22,850.1	$22,692.2	$22,611.2	$22,537.3

Figure 9. The market value of the Indonesian ICT sector 2017 – 2022 [29]

Moreover, being driven by the COVID-19 pandemic, which is a factor in accelerating changes in the ICT world, the telecommunications sector must start moving from the legacy telecommunications business (Voice, SMS) to the digital business. Figure 10 shows a graph of the Digital Telco Indonesia market specifically, which is not much different from the explanation in the previous section. The declining connectivity-legacy business has forced companies to move to-wards other businesses, namely data-connectivity, enterprise ICT, and digital life & smart platforms (Telkom Indonesia, 2020). It is expected that some of these business service products can increase CAGR up to 11-12%.
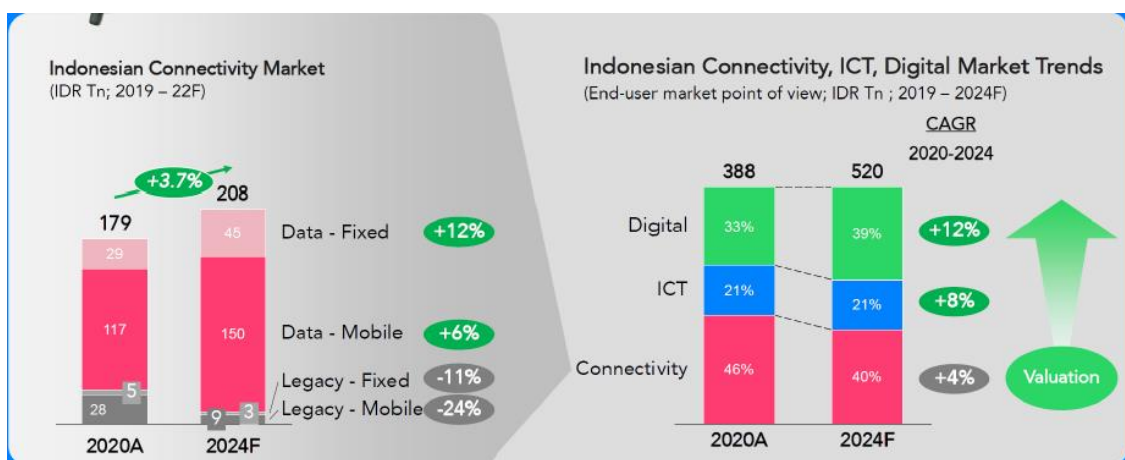

Figure 10. Indonesian digital telco market 2019 – 2025 [30]

In addition, the telecommunications company also needs to think about ways to enter a targeted and strategic digital business stream. The business stream is divided into dependent parts, each with its focus on product investment to generate maximum profitability. Figure 11 shows that three parts focus on digital business attention: Digital Connectivity, Digital Platforms, and Digital Services. 5G technology as an enabler of aspects of digital business is expected to be a new business opening for other digital platforms and services.
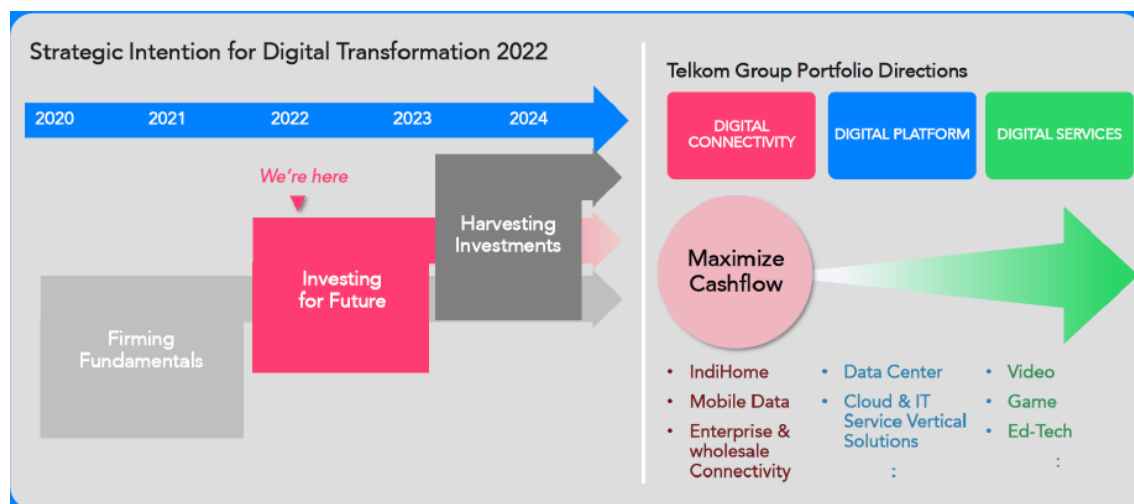
Figure 11. Strategic Intention for Digital Transformation [30]

Summary of data analysis result for these aspects: traffic, finance, network deployment, and digital services:

1. Traffic PT. MNO
   Total Cellular Connection tends to stagnant at 178 million connections in Q4-2021 and has not increased compared to connections 4 years ago (181 million connections in Q2-2017). ARPU has not increased since 6 years ago, namely from 13.01 USD in Q4-2015 to 3.01 USD in Q4-2021.

2. Financial PT. MNO
   Revenue has increased in the last 3 years, namely from 2.1 billion USD in Q2-2018 to 2.5 billion USD in Q3-2021.
   Capex/Revenue Ratio has been relatively stable (between 13% - 14%) in the last 2 years, namely 15.33% in Q2-2019 and 15.81% in Q3-2021.
   Net Profit was at the level of less than 500 Million USD in the last 2 years, Net Profit reached 507 million USD in Q4-2018 after that it was at the level of 4xx Millions USD in-cluding the latest data in Q3-2021 Net Profit of 467 million USD.

3. Network Deployment
   Network development in Indonesia is still uneven and there is a digital divide. Based on BPS data, the Top 5 Internet Access Distributions are the most consecutive in DKI Jakarta, D.I.Yogyakarta, Riau, East Kalimantan, and Bali. Meanwhile, The Bottom 5 Distribution of Internet Access is the lowest in Papua, NTT, North Maluku, West Sulawesi, and Maluku.

4. Digital Services
   Based on the ICT Market in Indonesia, Telecommunications revenue has experienced saturation, but Digital Services continues to show growth since the last 5 years (2017 until now).
   After Covid, the legacy telecommunications market fell rapidly, and the market moved from Legacy to Data Connectivity and Digital Business (for example Digital Life and Smart Plat-form rose by 11% with a market share of 18%).
   To maximize cash flow (profitability), PT. MNO makes Investments in the Digital Platform domain.

Based on those analyses (ARPU has not increased, Net Profit was less than 500 Million USD, Network development in Indonesia is still uneven and there is a digital divide), clearly shows that PT. MNO urges to do Business Transformation.

Conclusion

As one of the largest telecommunications companies in Indonesia, "PT. MNO Indonesia" should move from connectivity to a digital business. Some of the aspects of measuring business perfor-mance consist of traffic, finance, network development, and digital services. It is very im-portant to transform the telecommunications business from connectivity to digital business to sus-tain itself as the largest operator in Indonesia with a decent level of profitability. With the increas-ing interest in digital applications, the company should make investments and innovations in digital platforms that can create new business opportunities and maximize the monetization of digital ser-vices. The community's need for high-speed services increases rapidly, especially after Covid-19. Three interrelated digital business domains can be implemented: Digital Connectivity, Digital Plat-forms, and Digital Services.

References

[1]   Musnaini, A. Junita, H. Wijoyo, I. Indrawan, *Digital Business*. Banyumas: CV. Pena Persada, 2020.

[2]   Qualcomm, *Everything You Need to Know about 5G*. California: Qualcomm, 2021. https://www.qualcomm.com/5g/what-is-5g

[3]   Ibrahim Rohman, The BI Excel, Fitch Solutions, Inc. and Fitch Solutions, Ltd., 2022

[4]   McKinsey & Company, *The Road to 5G: The Inevitable Growth of Infrastructure Cost | McKinsey*. New York: McKinsey & Company, 2018.

[5]   D. Marabissi, L. Mucchi, R. Fantacci, M. R. Spada, F. Massimiani, A. Fratini, G. Cau, J. Yunpeng, & L. Fedele, "A real case of implementation of the future 5G city," *MDPI*, vol. 11, pp. 4, 2018. https://doi.org/10.3390/fi11010004

[6]   Badan Pusat Statistik, *Statistik Telekomunikasi Indonesia 2020*. Jakarta: Badan Pusat Statistik, 2020.

[7]   Berita Satu Research, *Indonesia Digital Infrastructure Report*. Jakarta: Berita Satu Media Holdings, 2021. https://research.beritasatu.com/download/indonesia-digital-infrastructure-report-2021.

[8]   Kementerian Komunikasi dan Informatika Republik Indonesia, *Percepat Bangun BTS 7.904 Lokasi, Kominfo Terapkan Terobosan Pinjam Pakai Lahan*. Jakarta: Pressrelease.id, 2021.

[9]   M. Ruffini, "Multidimensional convergence in future 5G networks," *Journal of Lightwave Technology*", vol. 35, no. 3, 2017.

[10]   M. Agiwal, H. Kwon, S. Park, & H. Jin, "A survey on 4G-5G dual connectivity: Road to 5G implementation," *IEEE Access*, vol. 9, 2021.

[11]   A. Ghosh, A. Maeder, M. Baker, & D. Chandramouli, "5G evolution: A view on 5G cellular technology beyond 3GPP release 15," *IEEE Access*, vol. 7, 2019.

[12]   S. Henry, A. Alsohaily, & E. S. Sousa, "5G is real: Evaluating the compliance of the 3GPP 5G new radio system with the ITU IMT-2020 requirements," *IEEE Access*, vol 8, 2020.

[13]   N. Vo, T. Q. Duong, M. Guizani, & A. Kortun, "5G optimized caching and downlink resource sharing for smart cities," *IEEE Access*, vol. 6, 2018.

[14]   M. Suryanegara, "Managing 5G technology: Using quality of experience (QoE) to identify the innovation enhancement pattern according to the indonesian market," *IEEE Access*, vol. 8, 2020.

[15]   W. S. H. M. W. Ahmad, N. A. M. Radzi, F. S. Samidi, F. Abdullah, M. Z. Jamaludin, & M. N. Zakaria, "5G technology: towards dynamic spectrum sharing using cognitive radio networks," *IEEE Access*, vol. 8, 2020.

[16]   N. N. Hikmah, "Mengelola infrastruktur bisnis digital," 2021. Accessible at: osf.io/g5h62.

[17]   L. Cano, G. Carello, M. Cesana, M. Passacantando, & B. Sanso, "Modeling the techno-economic interactions of infrastructure and service providers in 5G networks with a multi-leader-follower game," *IEEE Access*, vol. 7, 2019, 7: 162913-162940.

[18] I. A. Gedel, "A techno-economic comparison of 5G centralized wireless network architectures," *Przegląd Elektrotechniczny*, vol. 1, no. 5, pp. 14-21, 2021.

[19] E. J. Oughton & A. Jha, "Supportive 5G infrastructure policies are essential for universal 6G: Assessment using an open-source techno-economic simulation model utilizing remote sensing," *IEEE Access*, 2021.

[20] K. Sari, N. Nurliza, M. F. N. Fatimah, V. Melvia, & A. M. Putri, "Analisis lingkungan eksternal dalam menghadapi persaingan bisnis pada PT. Telkom Indonesia Tbk," *Jurnal Pendidikan Tambusai*, vol. 5, no. 2, pp. 3051-3056, 2021.

[21] S. D. Sam & S. Kanade, "telecom industry before and after reliance jio, "*International Journal of Social Science & Economic Research*, vol. 6, no. 4, 021.

[22] A. Tudzarov & S. Gelev. "Requirements for next generation business transformation and their implementation in 5G architecture." *International Journal of Computer Applications*, vol. 162, no. 2, pp. 31-35, 2017.

[23] A. Psyrris, A. Kargas, & D. Varoutas, "MNOs business models and roles enabled by 5G technologies and use cases: Transformation, challenges and strategies", in *2021 14th CMI International Conference - Critical ICT Infrastructures and Platforms (CMI),* 2021.

[24] S. K. Rao & R. Prasad, "Impact of 5G technologies on industry 4.0," *Wireless Personal Communication*, vol. 100, pp. 145-159, 2018.

[25] V. L. Shirokov, "Prerequisites and methodology for digital transformation of 4G networks into 5G ecosystem", in *International Conference on Distributed Computer and Communication Networks*, 2020, pp. 158-168.

[26] A. Manzalini, C. Buyukkoc, P. Chemouil, S. Kuklinski, F. Callegati, A. Galis, M. P. Odini, C. I, N. Crespi, E. Healy, & S. Sharrock, "Towards 5G software-defined ecosystems", *IEEE Software Defined Networks*, 2016.

[27] ITU, *Digital Infrastructure Policy and Regulation in the Asia-Pacific Region*. September 2019, Accessible at: https://www.itu.int/en/ITU-D/RegionalPresence/AsiaPacific/SiteAssets/Pages/Events/2019/RRITP2019/ASP/ITU_2019_Digital_Infrastructure

[28] S. Ariyanti, A. S. Slamet, & J. M. Munandar, "Study of mobile operator readiness measurement in indonesia for 5G technology deployment," *Buletin Pos dan Telekomunikasi*, vol. 9, no. 2, pp. 105-118, 2021.

[29] Frost & Sullivan, *Digital Market Overview: Indonesia*, Texas: A Frost & Sullivan White Paper (www.frost.com). 2018.

[30] Telkom Indonesia, *FU NITS Master Plan 2020-2025*, Jakarta: Telkom Indonesia. 2022.

Asah skill: An approach skill-based learning platform with management and strategic planning process

Ellin Devihana Pratiwi [1]*, Hendrik Amrico [2], Muhammad Afiff Rizki [3], Resti Savira [4], Catur Apriono [5]*

[1,2,3,4,5] *Department of Electrical Engineering, Faculty of Engineering, Universitas Indonesia*

*Corresponding Author: *catur.apriono@ui.ac.id*

Abstract: Digital transformation in education improves entities by implementing significant changes in education by combining information technology, computing, communication, and coordination. This paper proposes improving education through a digital transformation through platforms, specifically online skills-based learning called ASAH Skill. The survey was conducted by collecting question data from Google Forms. By using the Problem Validation method and Primary Data, obtained Product Validation on Customer Needs (Survey) and Product Strategy (Affinity Diagram). It is a market-based validation as a guide for designing ASAH Skill products that meet the criteria and needs of Skill Providers and Seekers. The method considered consists of three steps, namely strategic situation analysis (SSA), strategy formulation (SF), and strategy implementation (SI). We have finally found what customers need to develop their skills during the isolation period of the covid-19 pandemic. The survey was conducted on 51 respondents. Strategic Implementation consists of Mockup, Business Model Canvas (BMC), and Problem Solution Fit Canvas. The biggest detention experienced by respondents was very limited mobility (82.4%). The survey also showed that most of the activities respondents did during the pandemic were exploring new things (37.3%). Video editing is the most learned skill during the pandemic (25.9%), and YouTube is the respondents' main choice as a medium for learning skills (43.1%). Overall, the results of the study show that dispensing and learning skills are a new need for society during a pandemic. The provision of a platform is a development opportunity as a forum for the dissemination and learning of skills. The highest priority is on the 'easy to use' attribute where all respondents give an importance rating of 5.

Keywords: platform, online learning, skill, mockup

How to Cite: E. D. Pratiwi, H. Amrico, M. A. Rizki, R. Savira, & C. Apriono, "Asah skill : an approach skill-based learning platform with management and strategic planning process," *Matrix: Jurnal Manajemen Teknologi dan Informatika*, vol. 12, no. 2, pp. 91-101, 2022.

Introduction

The modern world is constantly changing. The innovation introduces various human activity fields so that people can continue developing, improving knowledge, skills, professions, or as simple as trying new types of activities. In addition, it requires creativity, willingness to work with colleagues to find new solutions, and the ability to critically develop the presentation of information, reliability, and logical intervention. Digitalization also changes the pattern of people's social life; it leads to open new opportunities to acquire and increase knowledge. Digital technology in the modern world is not only a tool, but also an existential environment that opens up new opportunities: learning anytime, lifelong learning, the ability to design educational journeys, and personal education [1].

The development of digital technology affects all aspects of society, especially in developing certain abilities (skills). The skill itself is the ability to act with a determined result with good implementation. Skills consist of two categories that are hard skills and soft skills. Hard skills, also known as technical skills, are abilities related to certain tasks or situations. These skills involve understanding and proficiency in certain activities that involve methods, processes, procedures, or techniques. These skills are easy to measure, unlike soft skills related to a person's personality.

In the current era of the industrial revolution on digital technology, Indonesia's number of internet users has increased. In March 2021, the number reached 212.2 million, or 77.9% of the total population of Indonesia [2]. This increment is mainly due to the COVID-19 pandemic, where people tend to spend time in quarantine to learn, explore certain skills, and kill boredom. It can also provide benefits for self-development. The fact is skill development (HR) in Indonesia is still very low, where the index value of Indonesian HR 2020 is 0.53 /HCI<0.60 [3]. This condition is partly due to the lack of public awareness regarding the importance of developing skills, both hard skills and soft skills.

On the other hand, COVID-19 affects the availability of job opportunities and the termination of employment contracts for most private workers. It increases the unemployment rate and reduces the value of community productivity. Economic inequality also occurs simultaneously with the increase in the unemployment rate. Data in August 2020 showed that the national employment opportunity rate decreased by 1.84% from the previous year's same period [4]. It also encourages the government to target 212,000 people to participate in the New Entrepreneurial Creation Program. The important training programs and skills development improve the quality and index of Human Resources in Indonesia, including high unemployment, scattered skill-providing media, and difficulty getting a job.

Based on previous research, COVID-19 has a significant impact on digital transformation in Hungary education, so new insights are necessary for digital transformation [5]. The industry 4.0 era encourages productivity, economic opportunities, and future jobs, thereby encouraging digital disruption [6], and could enhance a platform-based application business model [7]. The success of a digital learning platform depends on the user (user context and cultural context), social (environment), institutional background, and several other aspects [8].

The various skill-based platform has been researched but only specific to a certain skill such as entrepreneurship promotion [9], software skill learning [10], and researching skill [11]. This specified skill-based platform only benefit certain people who needed a specified skill, hence it is needed to have a new different skill-based platform to cover various kind of skills.

This paper proposes a product design platform based on a skill-based digital platform: ASAH Skill. This product design aims to provide a media (platform) that brings together skill providers and seekers in the informal sector (Hard Skills/Soft Skills). The word 'Asah' comes from an Indonesian language meaning 'sharpen'. For those who are training providers, ASAH SKILL can be a medium to create new jobs by utilizing the skills they have. As for skill seekers, they can take advantage to acquire the required skills. Problem Validation and Primary Data methods approach is used, Product Validation is obtained in Customer Needs (Survey) and Product Strategy (Affinity Diagram). It is market-based validation as a guide for designing ASAH Skill products that meet the criteria and needs of Skill Providers and Seekers. With ASAH Skill, it is to be expected to increase the national employment rate.

Methodology

The methodology will be used to ensure the approach of the skill-based learning platform can meet and be accepted by the market consists of three steps, which are strategic situation analysis (SSA), strategic formulation (SF), and strategic implementation (SI).
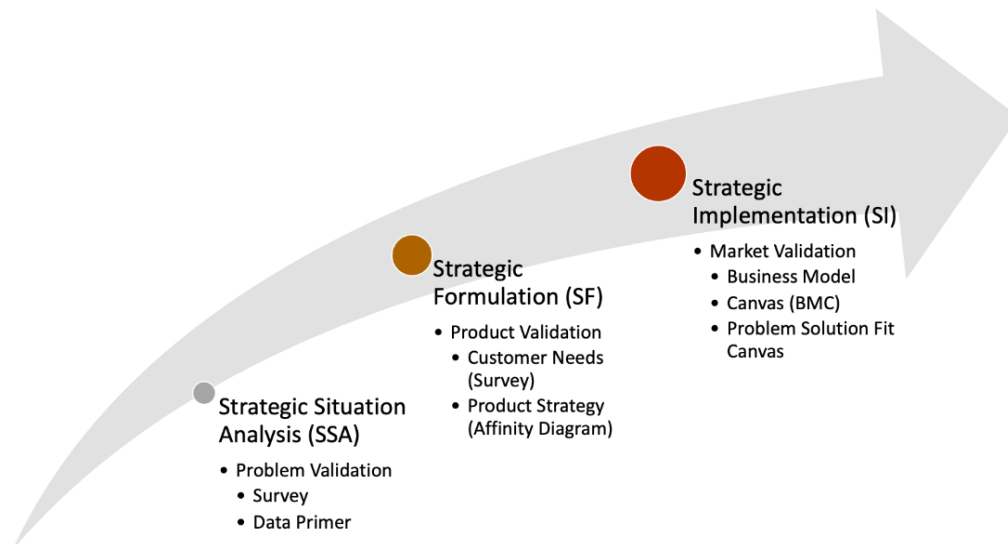
Figure 1. Methodology diagram
(Source: Data Collection by Researchers)

Strategic Situation Analysis (SSA)

Strategic situation analysis discovers the current condition to find the gap between ideal and existing conditions [12]. This step is to identify the problems in society. It is important to discover the needs of the society for a new form to upgrade their skills during the pandemic Covid-19 that brought new normal, limiting mobilizations, social interactions, and even work termination. SSA uses a survey to map customer behavior and needs with a simple questionnaire.

This research considers ten simple questions for the survey to determine a set of situations people face. The questions are to determine the detention that respondent is experiencing, what they usually do during the pandemic, what skills they want to learn, and what kind of media they use for learning skills.

Strategic Formulation (SF)

The strategic formulation uses available knowledge to document the intended direction of a business and the actionable steps to reach its goals [12]. In this strategic formulation step, we will focus on the product itself, where we analyze what kind of a platform-based product needs from customers. In this case, the platform is in the form of an app on a smartphone. This paper considers the form of a smartphone app for a couple of reasons. First, more people use smartphones rather than desktops for daily internet needs. Second, creating a mobile app design is simple and compact. Third, a mobile-based platform can easily transform into a web-based platform. The considered product validation consists of a customer's need and a product strategy.

This research surveys participants' demographics to find what customers need. The survey consists of 5 different dimensions: performance, reliability, feature, aesthetics, and serviceability. These dimensions are based on [13] and will turn into 20 questions, where each question corresponds to a different attribute. A scale of 1 to 5 for each question will determine how important an attribute is, where a score of 1 is not important, while a score of 5 is very important. After obtaining the survey, each attribute considers an average score.

The next step is a product strategy. This step will be conducted based on the survey's customer needs. The product strategy solutions will be used in the form of an affinity diagram analysis. An affinity diagram is a kind of predictive analysis technique that does the process of data collecting and fetches the insightful hiding correlation in between the dataset.

Strategic Implementation (SI)

Based on the SSA and SF, a strategic implementation refers to the process of executing plans and strategies [9]. SSA validates the problem, and SF validates the product. And defines the market itself. It is important to prepare the product to be accepted by the market. The market validation consists of Business Model Canvas (BMC) & Problem-Solution Fit Canvas.

The BMC is a visual way of identifying key elements of business and how they relate. Variables used on BMC are the problem, solution, key metrics, unique value proposition, unfair advantage, channels, customer segments, cost structure, and revenue stream. Go To Market Strategy will be generated from BMC. Meanwhile, Problem-Solution Fit Canvas is trying to validate BMC by focus group discussion with the user to get the user's perspectives.

Results and Discussions

SSA Results

Figure 2 shows the survey to determine what customers need to develop their skills during pandemic covid-19 isolation. The survey was conducted on 51 respondents. The survey was conducted by collecting question data from Google Forms. Most of the respondents are fresh graduates and adults with an age range of 24-30 years. The graphic below shows the result of the survey of situation analysis.
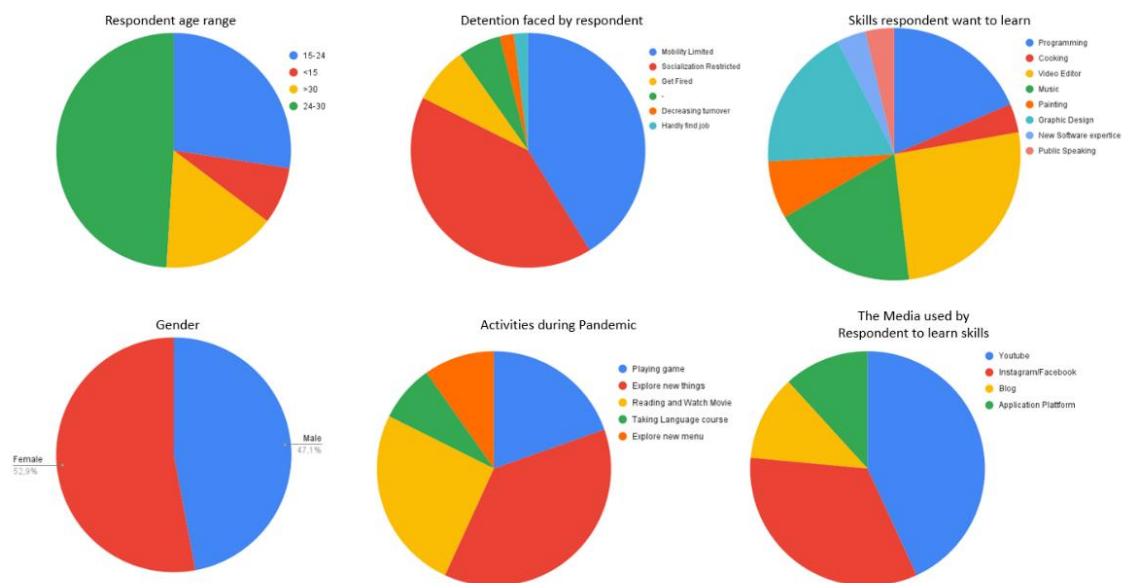
Figure 2. Survey of problem validation

The chart shows that most respondents are female, the age range at most is 24-30 years, and during the pandemic, the biggest detention experienced by respondents was very limited mobility (82,4%). The survey also demography that the most activities respondent do during a pandemic are exploring new things (37,3%). Video editors are the most popular skills to learn during a pandemic (25,9%), and YouTube became the main choice for responders as a medium to learn skills (43,1%). Overall, the result shows that dispensing and learning skills are a new need for society amid a pandemic. The provision of a platform is an opportunity for development as a forum for dispensing and learning skills.

SF Result

Strategic Formulation consists of Customer Needs and Product Strategy. Strategic Formulation consists of Customer Needs and Product Strategy. Consumer needs are known by spreading questions according to consumer needs for the product. The product strategy is made to determine the main priorities of consumer wants and needs for platform products. With the form of an affinity diagram, it can be seen that the highest consumer interest in the platform in terms of what must exist and which can support the use of the platform.

Customer Needs (Survey)

Based on the 20 questions related to customer needs, it shows that 50 respondents found the main attribute the skill-based platform needs. In this case, the platform is in the form of a smartphone app. These respondents are the same as the SSA survey given.

Table 1. Power consumption, cost and time

Dimension	Attribute	Importance Scale (Average)
Performance	The application recognizes user preference	4.28
Performance	The application can give important notification	3.66
Performance	The application provides any skills	4.32
Performance	Skill-seeker can find various skills in the app	4.68
Performance	Show skill provider profile	4.82
Reliability	Easy to use	4.9
Reliability	Low-cost internet access	3.96
Reliability	Less memory capacity	3.96
Reliability	Application runs without error	4.84
Reliability	The application can run on any OS	4.58
Reliability	A synchronize-able account with another platform (Google, FB)	3.66
Reliability	Payment can be done by application	4.14
Reliability	The application can be used anywhere and anytime	4.74
Feature	Have 'send feedback' feature when an error occurred	3.88
Feature	Application categorizes the skills	4.68
Feature	Have a "chat" feature between provider and seeker	4.6
Feature	Certificate at the end of the course	4.22
Feature	How to use app tutorials	4.4
Aesthetics	Attractive interface	4.38
Service ability	Application has customer service	4.02

Based on Table 1, the average score on the importance scale is 4.34. A score above 4.43 is considered the main attribute, meaning the product must include those attributes. The main attribute can be seen in Table 1 with bolded text. From table 1, we can see that the highest priority is within the 'easy to use' attribute, where all respondents gave an importance rating of 4.9.

Product Strategy (Affinity Diagram Analysis)

The product strategy will be in the form of an affinity diagram analysis. Affinity diagrams are similar to cause and effect diagrams but are more suitable for larger issues where even the main issue is not well defined [14]. The main attributes can then be modified from the customer needs into a solution of the product itself. Based on Table 1, we can create a diagram consisting of four different technical, account, features, and customer service categories. The detail of each category of the affinity diagram analysis can be seen in Figure 3.
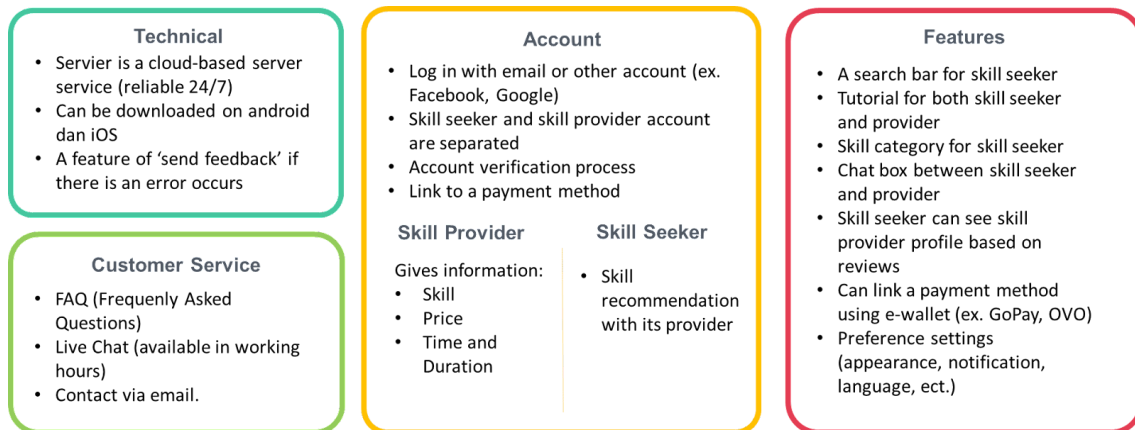
Technical
- Servier is a cloud-based server service (reliable 24/7)
- Can be downloaded on android dan iOS
- A feature of 'send feedback' if there is an error occurs

Customer Service
- FAQ (Frequently Asked Questions)
- Live Chat (available in working hours)
- Contact via email.

Account
- Log in with email or other account (ex. Facebook, Google)
- Skill seeker and skill provider account are separated
- Account verification process
- Link to a payment method

Skill Provider

Gives information:
- Skill
- Price
- Time and Duration

Skill Seeker
- Skill recommendation with its provider

Features
- A search bar for skill seeker
- Tutorial for both skill seeker and provider
- Skill category for skill seeker
- Chat box between skill seeker and provider
- Skill seeker can see skill provider profile based on reviews
- Can link a payment method using e-wallet (ex. GoPay, OVO)
- Preference settings (appearance, notification, language, ect.)

Figure 3. Affinity diagram analysis

Technical

The technical category is created because of the high priority of a platform to run smoothly. The 'Application runs without error' and 'Application can be used anywhere and anytime' attributes can be taken into account in the form of a cloud server-based service. The cloud-based server is usually reliable, where it can be guaranteed to run 24/7 without internal errors. This category includes the platform app that can run into different operating systems and error handling in the form of 'send feedback'.

Account

The account category is made for skill seeker and skill provider accounts. From the attribute 'show skill provider profile' and 'have a "chat" feature between provider and seeker,' we separated the skill seeker and skill provider accounts. It will help the platform to differentiate the roles. The skill provider needs to have a clear profile of the skill seeker, so the information needed for the skill provider are skill information, price, time and duration, experiences, etc. Both skill seeker and provider also need to verify their account. It is needed to validate whether the account is a real person.

Features

The features category covers all the features needed for the platform app. The 'skill-seeker can find vary skill in app' attribute can be in the form of a search bar, and the 'how to use app tutorial' attribute can be in the form of a tutorial for both skill seeker and provider. This category also consists of a category for skill, a chat box between skill seeker and provider, profile viewing for skill provider based on reviews, payment method feature, and preference settings such as appearance, notification, language, etc.

Customer Service

The customer service category consists of frequently asked questions, live chat for customer services, and contact information. This category is not as important as the other three categories, but it is needed to maintain the platform's reliability in assisting customers.

SI Results

The strategic implementation consists of Mockup, Business Model Canvas (BMC), and Problem Solution Fit Canvas.

Mockup

A mockup design is made based on the affinity diagram analysis. With the help of the affinity diagram, a clear idea of what the platform might look like in the form of a smartphone app.
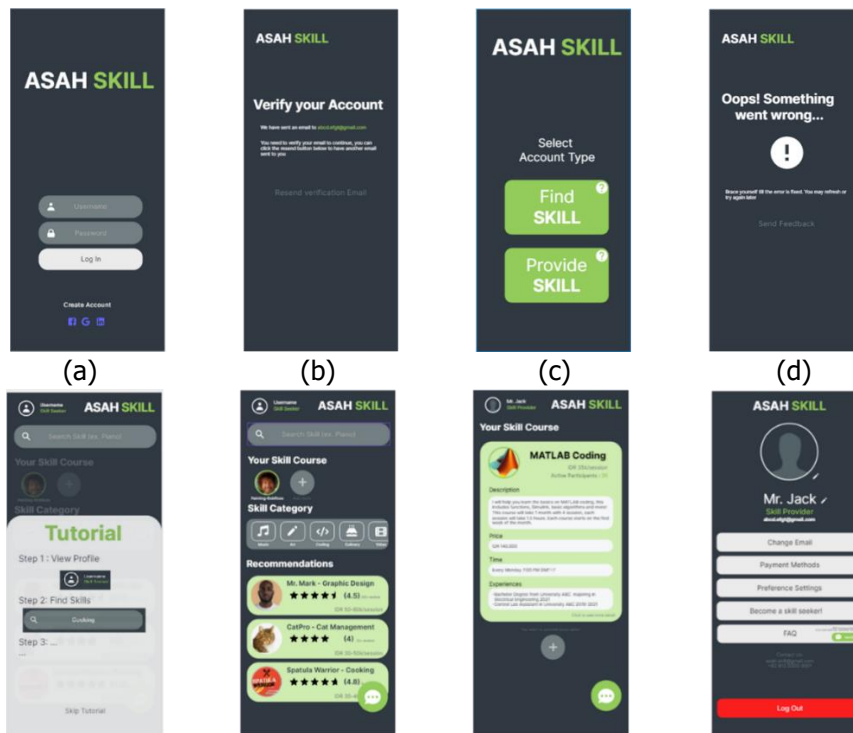
(a)   (b)   (c)   (d)

Figure 4. Mockup design: (a) login panel; (b) account verification; (c) separate account; (d) error message; (e) tutorial panel; (f) skill seeker main panel; (g) skill provider main panel; (h) ac-count panel

The mockup design from Figure 4 represents the key attribute needed for the application platform. Based on the affinity diagram analysis, six different panels are shown.

Business Model Canvas (BMC)

Business Model Canvas (BMC) helps us map out the marketing strategy to get the market's overview [15]. BMC can develop a clear view of your problem, solution, key metrics, value proposition, unfair advantage, channels, customer segments, cost structure, and revenue stream. The detail of every element in BMC will represent in Figure 5.

Figure 5. Business Model Canvas (BMC)

Problem

Problems on BMC is defining the problems behind this idea. ASAH Skill is designed to overcome the problems of the high unemployment rate, but the media of skill-based learning is still scattered. People find it difficult to find work. There is no platform for providing informal skills. There is no media for skill owners to be productive from their skills, and the trend of learning new skills with free time during quarantine.

Based on our research, the existing solution that has already come up for those problems is a platform that only provides formal education and skill-based learning held by the studio on the website/social media.

Solution

Our proposed solution that we provide to overcome the problems is providing a platform that brings together skill providers with skill seekers and provides a platform that provides formal skills and informal skills.

Key Metrics

The key metric is being used to measure the success rate of our solution. ASAH Skill's key metrics are the number of users (skill provider & skill seeker) and income earned by a fee on transactions on the platform and advertisement.

Unique Value Proposition

The unique value proposition is being used to show the product's highlight and how it differs from the existing product that has already been launched on the market. We give two perspectives of our product's value proposition that is Skill Seekers, who are looking for skills that are easier, varied, and affordable without being limited by space & time, and Skill Providers, who provide a platform to channel skills, so they can generate and give benefit to many people. On the high-level concept, ASAH Skill is a platform where people can get skills easily, varied, and at affordable prices without being limited by space & time.

Unfair Advantage

Unfair advantage is being used to show the benefits that users get. ASAH Skill's unfair advantages are a collaboration with the community, a platform that focuses on providing informal skills (hard skills / soft skills), granting of certification issued by an official organization, providing exposure for users to recruiters through personas, and providing a "gallery" feature for users to share the work of their skills.

Channels

ASAH Skill's channels are community, social media (YouTube, FB, IG, Twitter, LinkedIn, TikTok), content distribution platforms, word of mouth, user referrals, and partnerships (Influencers & KOL).

Customer Segments

In this element, we are trying to focus on our product's customer segment. It will help the product be accepted by market accept the product. Our customer segments are graduated middle/high school, productive workers (24-30 years old), and SES ABC. Meanwhile, our specific customer for early adopter is college student, workers (22-27 years), and people who lost their jobs due to the pandemic in Jakarta.

Cost Structure

ASAH Skill's allocated cost is being used for the development platform, marketing in all channels defined above, and human resources to run the company.

Revenue Stream

In this element, we are trying to find the benefit that the company gets by build/investing in this product. We have four streams that charge 10% of the trainer rate based on class/session, subscriptions (bundling classes/sessions), advertising, and add-ons (extra time, tools, etc.).

Problem Solution Fit Canvas

Problem Solution Mapping Canvas is not only a mapping but also an enabling translation where you turn problems into solutions and communication strategies, taking customer behavior into account to increase the chances of receiving solutions. It gives us a glimpse of ideas that might fit into reality. In this paper, the problem fit canvas is used as BMC validation to answer the purpose of the ASAH skill product design. The Problem Solution Fit Canvas can be seen in Figure 6.
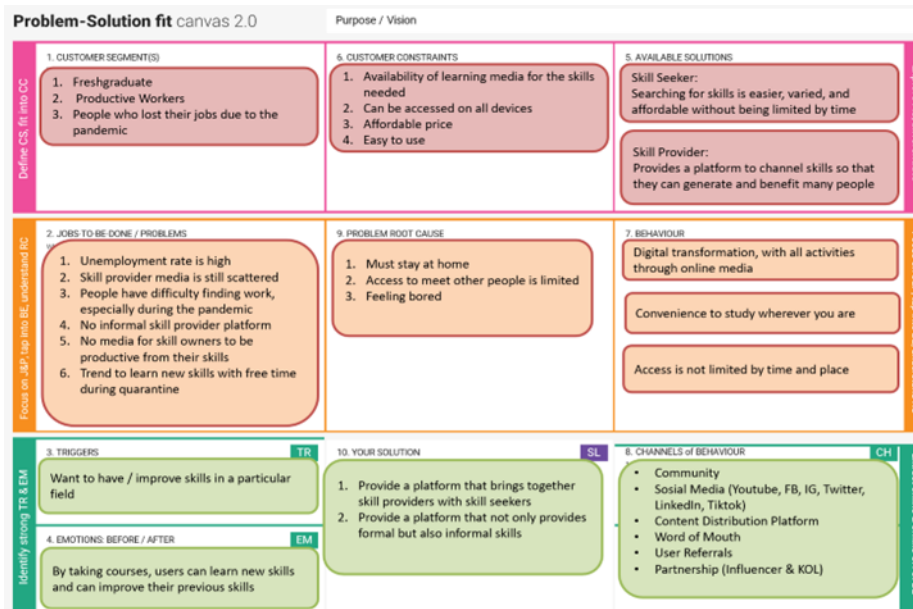

Figure 6. Problem Solution Fit Canvas
(Source: Data Collection by Researchers)

Customer Segment

Customer segment is used to determine the target market of the product. People from the younger generation can and will accept and use the product as it functions and uses it. In the ASAH SKILL product design, the target market for the product is young people who have just graduated from school (fresh graduates), productive workers, and people who have lost their jobs due to the COVID-19 pandemic.

Problems

Problems describe the problems behind which the product is designed and made. ASAH SKILL is designed to overcome problems including the high unemployment rate, the availability of skill providers that are still scattered, the difficulty of finding work, the trend of learning new skills, the absence of an informal skill provider platform, the absence of media for skill owners to be productive from existing skills.

Triggers

Triggers are things that trigger people's desire to use the ASAH SKILL product design. The desire to have or improve new skills in certain fields is the main trigger why customers need this media platform.

Emotion

Emotion aims to ensure that the product design goals can be achieved and follow users' wishes. So that by taking courses, users can learn new skills and improve their previous skills.

Customer Constraints

Customer constraints aim to determine user wants and needs that must be available on a skill development media platform. In this case, customer needs for the platform include the

availability of learning media for the skills needed, can be accessed on all devices, affordable prices, and easy to use.

Problem Root Cause

Describing the background of the problem emerged the concept of designing the ASAH SKILL media platform. The problems are due to the pandemic, community activities are limited, people are not encouraged to do activities outside the home, access to meet other people is limited, and there is a sense of boredom and boredom because they don't have activities.

Available Solution

Available solution is a solution that can be provided for users of the platform design, both skill seekers and skill providers. Skill seekers will get solutions to find skills easier, varied, and at affordable prices without being limited by space and time. On the other hand, skill providers can provide a platform to channel skills so that they can generate and benefit many people

Your Solution

It is suggested that we provide to answer the needs of users (searchers and skill providers) by providing a platform that brings together skill providers with skill seekers and provides a platform that provides not only formal skills but also informal

Behavior

Habits of people in the current digitalization era use online media to make work easier, get the convenience of studying and working anywhere, and have easy access without limits of space and time.

Channel for Behavior

These are channels that users commonly use to meet their needs. The channels include community, social media (YouTube, FB, IG, Twitter, LinkedIn, TikTok), content distribution platforms, word of mouth, user referrals, and partnerships (Influencers & KDL).

Conclusion

The need for skills is found in the productive age range who experience mobility barriers during the COVID-19 pandemic. Based on the data, 43,1% of the medium used to learn and channel skills are YouTube. But 82.4% of respondents chose there are some obstacles such as less interactive & scattered information available. From the survey results, it can also be seen that learning skills through a specific channel are a necessity. So that this can be an opportunity in providing an application-based platform that can improve the quality of learning from users.

From above, we got the opportunity to provide an application-based platform. To fulfill the customer needs, we need to survey on a scale of 1 to 5 to see which parameters were the main focus of the application. The scale of the greatest interest lies in the ease of application (easy to use) and the provision of profiles for skill providers. Based on the data of customer needs, we formulate the product strategy using affinity diagrams for ASAH SKILL that can be categorized into Technical, Customer Service, Account, and Features. The technical category is created because of the high priority of a platform to run smoothly. Customer Service is needed to maintain the platform's reliability in assisting customers. The account category is made for skill seeker and skill provider accounts. The Features category consists of a category for skill, a chat box between skill seeker and provider, profile viewing for skill provider based on reviews, payment method feature, and preference settings such as appearance, notification, language, etc.

In the last step, we get insight on how to communicate business ideas or concepts to be more accepted by the market through clear segmentation, our specific customers for early adopters are college students, workers (22-27 years), and people who lost their jobs due to the pandemic in Jakarta. Channels used, are channels that users commonly use to meet their needs including community, social media (YouTube, FB, IG, Twitter, LinkedIn, TikTok), content distribution platforms, word of mouth, user referrals, and partnerships (Influencers & KDL). The distinguishing value proposition with existing platforms where people can get skills easily, varied,

and affordable prices without being constrained by space and time. It is also supported by Problem Solution-Fit Canvas which can validate proposed ideas for customer needs that have not been offered by other applications.

Acknowledgments

Intended to pay gratitude to sponsors, fund bearers, resource persons, and other parties that play pivotal roles in the study. The author needs to ask for permission from persons or institutions for mentioning them in the acknowledgments. Editors need not be acknowledged.

References

[1]   A. A. Bilyalova, D. A. Salimova, & T. I. Zelenina. "Digital Transformation in Education," In *Lecture Notes in Networks and Systems*, 265–76, 2020.

[2]   C. M. Annur. *Ada 2047 juta pengguna internet di Indonesia awal 2022*. Jakarta: Databoks, 2022. Available online: https://databoks.katadata.co.id/datapublish/2022/03/23/ada-2047-juta-pengguna-internet-di-indonesia-awal-2022# (accessed on 25 April 2022).

[3]   World Bank. "The Human Capital Index 2020 Update," *The Human Capital Index 2020 Update*, 2020. https://doi.org/10.1596/34432

[4]   M. A. Rizaty. *Tingkat Kesempatan Kerja DKI Jakarta Terendah Nasional.* Jakarta: Databoks, 2021. https://databoks.katadata.co.id/datapublish/2021/02/18/tingkat-kesempatan-kerja-dki-jakarta-terendah-nasional (April 20, 2022).

[5]   Bogdandy, Bence, J. Tamas, & Z. Toth. 2020. "Digital Transformation in Education during COVID-19: A Case Study," *11th IEEE International Conference on Cognitive Infocommunications, CogInfoCom 2020 - Proceedings*, , pp 173–78, 2020.

[6]   L. Caruso. "Digital Innovation and the Fourth Industrial Revolution: Epochal Social Changes?," *AI and Society,* vol. 33, no. 3, pp. 379–392, 2018.

[7]   B. Berlilana, R. Utami, & W. M. Baihaqi. "Pengaruh Teknologi Informasi Revolusi Industri 4.0 terhadap Perkembangan UMKM Sektor Industri Pengolahan," *Matrix : Jurnal Manajemen Teknologi dan Informatika*, vol. 10, no. 3, pp. 87-93, 2020.

[8]   N. Mehta, S. Chauhan, P. Gupta, & M. P. Jaiswal. "Pursuing digital learning platform success: A meta-analytic investigation of user and cultural contingencies," *Communications of the Association for Information Systems*, vol. 48, no. 1, pp 34, 2021.

[9]   H. Ejo-Orusa, & D. L. Mpi. "Reinventing the 'Nwaboi'apprenticeship system: A platform for entrepreneurship promotion in Nigeria," *International Journal of Advanced Research in Management and Social Sciences*, vol. 8, no. 9, pp. 98-130, 2019.

[10]   P. Rani. "Strategy Implementation in Organizations: A Conceptual Overview," *Management*, vol. 14, pp. 205-218, 2019. Doi: 10.26493/1854-4231.14.205-218.

[11]   G. Paulmani. "Improving research skills with lean-agile-based instruction: Lean-agile led, analytics oriented platform for flexible, self-regulated, inclusive learning," *Proceedings - IEEE 20th International Conference on Advanced Learning Technologies, ICALT 2020*, pp. 387–389, 2020. https://doi.org/10.1109/ICALT49669.2020.00122

[12]   Z. Svobodová, & J. Rajchlová. "Strategic behavior of e-commerce businesses in online industry of electronics from a customer perspective," *Administrative Sciences*, vol. 10, no. 4, 2020.

[13]   W. H. B. Munthe, A. A. Suhendra, & R. M. E. Hadi. "Analisis Kebutuhan Pelanggan Menggunakan Aplikasi Emma Smart Home Menggunakan Integrasi Product Quality dan Model Kano," *eProceedings of Engineering,* vol. 4, no. 2, 2017.

[14]   C. Cote. *5 Tips for Successful Strategy Formulation | HBS Online*. England: Harvard Business School, 2020. (n.d.). Retrieved May 10, 2022, from https://online.hbs.edu/blog/post/strategy-formulation

[15]   Sheda. *How To: Business Model Canvas Explained.* Retrieved May 10, 2022, from https://medium.com/seed-digital/how-to-business-model-canvas-explained-ad3676b6fe4a

Application of the Blowfish Algorithm in securing patient data in the database

Pahrul Irfan [1], *Rifqi Hammad* [2]*, *Andi Sofyan Anas* [3], *Fatimatuzzahra* [4], *Nanang Samudra* [5]

[1,2,3,4,5] *Bumigora University, Indonesia*

*Corresponding Author: *rifqi.hammad@universitasbumigora.ac.id*

Abstract: Patient data is one of the datasets managed by the hospital. Patient data in the form of examination results and other data is important data that is private and confidential. Therefore, patient data needs to be secured so that there is no misuse of data by parties who are not responsible for things that can harm the data owner. One of the several methods that can be used to secure data is cryptography. Cryptography itself has several algorithms, one of which is blowfish. This study applies the blowfish algorithm to secure patient data in the database to reduce the possibility of data misuse by irresponsible parties. This study succeeded in implementing the blowfish algorithm for securing patient data. The data stored in the database is the result of encryption using the blowfish algorithm, the results of which are difficult to understand because there is a combination of symbols and text. The application of the algorithm affects the data storage time in the database, which originally took 0.12 seconds to save data and now takes 0.28 seconds to store data.

Keywords: data patient, cryptography, blowfish, database

How to Cite: P. Irfan, R. Hammad, A. S. Anas, Fatimatuzzahra, and N. Samudra, "Application of the Blowfish Algorithm in securing patient data in the database," Matrix: Jurnal Manajemen Teknologi dan Informatika, vol. 12, no. 2, pp. 102-108, 2022.

Introduction

The utilization of information technology in the current era has an important role in the success of an organization [1] because information technology can help all organizational activities such as administration and others [2]. Such is the case with organizations in the health sector that utilize information technology in providing services in the health sector [3]. The hospital is one of the health organizations that utilize information technology. In managing its data, it uses digital storage media in the form of a database [4].

The data managed by the hospital is in the form of patient data, examination results data, and others. Patient data is very important data, so it needs to be protected because it is private and confidential [5] [6]. This data can be stolen through system hacks carried out by irresponsible and misused elements, such as identity theft and filing false insurance claims [7]. This is like what happened in 2020 when Indonesian citizen data related to COVID-19 was stolen by unauthorized parties and then sold on the dark web via the RapidsForum forum [8].

To anticipate the misuse of data by unauthorized parties, it is necessary to provide security for the data. This security can be done in the form of data encryption so that the data is difficult to open or read [9]. The method that can be used to perform encryption is called cryptography [10]. Data cryptography is one of the techniques used to secure data by encrypting it [11] so that it can be turned into other random data [12]. Cryptography has several algorithms that are used in the data security process, such as the Caesar cipher, vigenere cipher [13], Blowfish, and others. Blowfish is one of the cryptographic algorithms used to encrypt and decrypt data that operates with the input and output of data blocks of varying sizes, ranging from 32 bits to 448 bits [14].

In this study, the algorithm method used to overcome the problem of misuse of patient data is the blowfish algorithm, in which the algorithm will be embedded in the information system used to manage patient data. so that the data stored in the database will be different from the data entered. This is because the data is encrypted. It is hoped that the application of the blowfish algorithm will reduce the misuse of patient data that can harm the data owner. To support the research carried out, there are several studies related to the security of patient data and the use of the blowfish algorithm as data security, such as that conducted by Erwin Gunandhi and Agung Sudrajat entitled "Securing Medical Record Data Using Vigenere Cipher Cryptography". This study resulted in a medical record data security system using the vigenere cipher method [15]. The difference between the research conducted by Erwin and this research is that this study uses the blowfish method as the algorithm used in encrypting patient health data.

Another research is a study conducted by Nuniek Fahriani and Indah Kurniawati entitled "Patient Data Security with the Blowfish Algorithm on HOTSPOTDT". This study resulted in an application that was used to encrypt medical record files from patients at the HOTSPOTDT hospital [16]. The difference between the research conducted by Nuniek and Indah and the research to be conducted is that the research to be conducted does not encrypt medical record files but encrypts the data entered by the user into the database to reduce the misuse of the data stored in the database.

Another research is a study conducted by Erick Erwin Nylis and Purwanto entitled "Implementation of Email Security Using the Web-Based Blowfish Method at UPT Puskesmas Pondok Kacang Timur". This study resulted in an application that was used to encrypt email messages related to technical services provided or carried out by the UPT Puskesmas Pondok Kacang Timur [17]. The difference with the research that will be carried out is that in the research to be carried out, encryption is carried out on the data entered by the user into the database, so that the data available in the database is encrypted and difficult to understand, while in the research conducted by Erick and Purwanto, encryption is carried out on e-mail messages.

In addition to the three studies previously mentioned, there are also studies related to data protection using blowfish conducted by Harlen Gilbert Simanullang and Arina Prima Silalahi with the title "Blowfish Algorithm To Improve Mysql Database Security". This research produces an application that can be used to perform encryption and decryption with the blowfish algorithm [18]. The difference with the research that will be carried out by Harlen and Arina with this research is that this research produces an additional feature that is embedded in the patient data management application where every data input process into the database, the data will be encrypted automatically where the data stored is data that has been encrypted

Methodology

This study uses the blowfish algorithm method for encrypting patient data. The flow of the application of blowfish in this study can be seen in Figure 1.
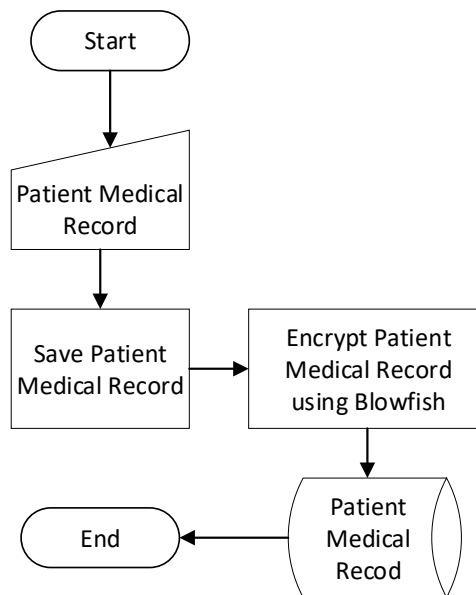
Figure 1. Blowfish Algorithm Implementation Flow

Figure 1 shows that the encryption process is carried out on the data before it is entered into the database. Blowfish is a cryptographic algorithm that is quite strong and has a large enough space and various lengths. Blowfish also keep their keys secret [19]. The encryption process using blowfish in this study can be seen in Figure 2.
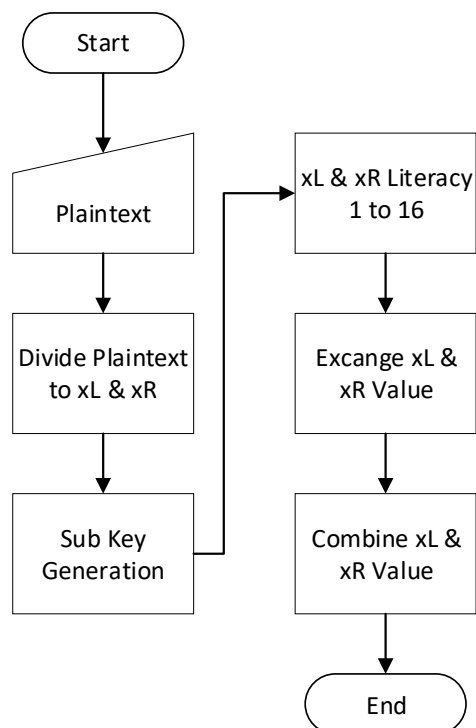

Figure 2. Encryption Flow

Figure 2 shows the encryption process used in the study with the blowfish algorithm. There are several steps involved in this process, namely:
1. The data to be stored in the database (plaintext) will be divided into 2, namely xL and xR
2. Key generation by creating an array P from P0 to P17 in hexadecimal form and converting it into binary form

3. Making S arrays, each of which is 255 in hexadecimal form and converted in the form of binary numbers, the plaintext is converted into binary form and divided into 2 (two) namely xL and xR, then enters the first round where i = 0 where xL is XoR right with P0 which was previously created in the P array, this xL is divided by 4 into 8 bits, namely a,b,c,d, then look for F(function) with equation (1) [20].

$$F(XL) = (((S0.a + S1.b \bmod 2^{32})XoR\ S2.c) + S3d\ mod\ 2^{32})$$ (1)

After finding the value of F, then next look for the value of xR. Finding the value of xR can be done by using the equation (2) [20]

$$XR = F(XL)XoR\ XR$$ (2)

After finding the xR value, the process of exchanging xL values with xR is carried out for up to 16 liters. This process generates new xL and xR values, each of which has 32 bits.
4. xL and xR values are exchanged again, then XoR values xL with P16 and xR with p17
5. Finally, combine the xL and xR values so that you get the ciphertext result

Results and Discussions

The encryption process with blowfish on patient data changes the shape of the data stored in the database. For example, the data entered into the system is shown in Figure 3.

Figure 3. Patient Data Form

In Figure 3, the patient data fields that will be stored are the doctor's name, patient name, date of birth, address, patient complaints, diagnoses, and actions. The fields will be encrypted using blowfish with the keyword "blowfish". The display on the database before the patient data is encrypted can be seen in Figure 4.

DOKTER	PASIEN	TANGGAL_LAHIR	ALAMAT	KELUHAN	DIAGNOSA	TINDAKAN
dr. Gede Supartha, Sp.M	Kurniadin Putra	06/14/2007	Mataram	Demam Sakit Tenggorakan Suara Serak	Radang Tenggorokan	Pemberian obat demam dan radang

Figure 4. Data Before Encryption

Figure 4 shows the results of data storage without encryption. These results indicate that the data entry can still be misused because it displays the true meaning. Therefore, encryption is necessary. The results of the encryption can be seen in Table 1.

Table 1. Encryption Result

Attribute Name	Plain Text	Encrypt Result
Nama Dokter	*dr. Gede Supartha, Sp.M*	c ë ¥ ¦ ö Ã z ú { Ñ $. ¨ . ® p © \ ï H ? Ü G Ï
Nama Pasien	*Kurniadin Putra*	. z Ý Ó & ¢ ð o . û Þ À « & T k
Tanggal Lahir	*06/14/2007*	H 4 [¦ ã j ¼ à d ¨ i < ¹ ¤ .
Alamat	*Mataram*	I . ¥ ï µ ¸ K .
Keluhan	*Demam* *Sakit Tenggorakan* *Suara Serak*	¬ œ Ê õ . 9 ç ê . ü ª ª s' . . ? • . · ñ l Á p G . Õ . < . Q … Ö . û » . ¤ . N
Diagnosa	*Radang Tenggorokan*	ã 9 ì Y É Õ Ë . B i . ¸ . p . º ë . • . º ê ã
Tindakan	*Pemberian obat demam dan radang*	` ó c . . q . . . l / ý C ù z Þ ª ± Ã , • c Ï . . Ì . T x Ê ê .

Table 1 shows the results of the encryption of the plaintext that will be stored in the database. So there is a change in the stored data, as can be seen in Figure 5.

Figure 5. Encryption Results in Database

Figure 5 shows the encryption results that have been stored in the database. The encryption process affects the speed of data storage. This can be seen from the time it takes to save data to the database. The time needed to store data without any encryption process is shown in Figure 6.
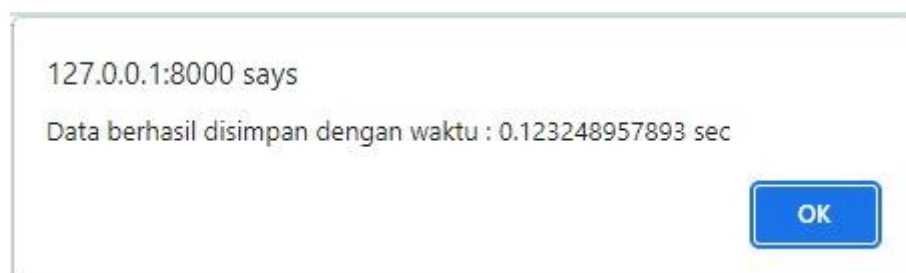

Figure 6. Time Required Before Encryption

Figure 6 shows the time required by the system to store data without any encryption process is 0.12 seconds. Meanwhile, the encryption process takes a little longer. This is shown in Figure 7.

Figure 7. Time Required After Encryption

Figure 7 shows the time required by the system to store data is 0.28 seconds. There is a difference of about 0.16 seconds between before and after encryption. This time difference occurs because of the encryption process carried out.

Conclusion

Based on the research conducted, the application of the blowfish algorithm in securing patient data can be carried out and produces ciphertext that is quite difficult to understand, but with the application of the algorithm, the data storage process becomes slightly longer than before, from 0.12 seconds to 0.28 seconds. As for suggestions for future research, it is to develop this blowfish algorithm again and it can also be combined with other cryptographic methods to produce ciphertext that is more difficult to crack so that the data remains safe and is not misused by unauthorized parties.

References

[1]   R. Hammad, A. C. Nurcahyo, A. Z. Amrullah, P. Irfan, and K. A. Latif, "Optimization of data integration using schema matching of linguistic-based and constraint-based in the university database," *J. Manaj. Teknol. dan Inform.*, vol. 11, no. 3, pp. 119–129, 2021.

[2]   Y. Farida and L. N. Desinaini, "Designing a microsoft access-based administration letters and archives system at BPJS of employment regional office of East Java," *Matrix J. Manaj. Teknol. dan Inform.*, vol. 11, no. 1, pp. 42–54, 2021.

[3]   A. M. Ningtyas and I. K. Lubis, "Literatur Review Permasalahan Privasi Pada Rekam Medis Elektronik," *Pseudocode*, vol. 5, no. 2, 2018, doi: 10.33369/pseudocode.5.2.12-17.

[4]   H. D. Siregar, F. S. Sulaiman, and N. Falih, "Literatur Review Permasalahan Pengamanan pada Database," in *Seminar Nasional Mahasiswa Ilmu Komputer dan Aplikasinya (SENAMIKA)*, 2021, pp. 521–530.

[5]   A. Riski, A. Kamsyakawuni, and M. Z. Ari, "Implementasi Vigenere Cipher Pada Pengamanan Data Medis," *J. Ris. dan Apl. Mat.*, vol. 2, no. 1, pp. 23–30, 2018.

[6]   N. O. Akande, C. O. Abikoye, M. O. Adebiyi, A. A. Kayode, A. A. Adegun, and R. O. Ogundokun, "Electronic Medical Information Encryption Using Modified Blowfish Algorithm," in *Computational Science and Its Applications – ICCSA*, 2019, pp. 166–178.

[7]   L. Sutandra, "Pengaruh Sistem Pengamanan Data Pasien di Rumah Sakit Menuju Era Revolusi Industri 4.0," *J. Heal. Sci. Physiother.*, vol. 1, no. 2, pp. 106–114, 2019.

[8]   C. Indonesia, "Deretan Peristiwa Kebocoran Data Warga RI Sejak Awal 2020," *CNN Indonesia*, 2020. https://www.cnnindonesia.com/teknologi/20200623160834-185-516532/deretan-peristiwa-kebocoran-data-warga-ri-sejak-awal-2020 (accessed May 05, 2022).

[9]   I. Riadi, A. Fadlil, and F. A. Tsani, "Pengamanan Citra Digital Berbasis Kriptografi Menggunakan Algoritma Vigenere Cipher," *J. Inform. Sunan Kalijaha*, vol. 7, no. 1, pp. 33–45, 2022.

[10]   D. Erdriani and D. Devita, "APLIKASI MATRIK PADA ILMU KRIPTOGRAFI DENGAN MENGGUNAKAN MATLAB," *J. KomtekInfo*, vol. 8, no. 2, pp. 154–162, 2021.

[11]   Hermansa, R. Umar, and A. Yudhana, "Analisis Sistem Keamanan Teknik Kriptografi dan Steganografi Pada Citra Digital (Bitmap)," in *Seminar Nasional Teknologi Fakultas Teknik Universitas Krisnadwipayana*, 2019, pp. 520–528.

[12]   A. Fadlil, I. Riadi, and A. Nugrahantoro, "Data Security for School Service Top-Up Transactions Based on AES Combination Blockchain Technology," *Lontar Komput. J. Ilm. Teknol. Inf.*, vol. 11, no. 3, p. 155, 2020.

[13]   V. C. Hardita and E. W. Sholeha, "Penerapan Kombinasi Metode Vigenere Cipher, Caesar Cipher dan Simbol Baca Dalam Mengamankan Pesan," *J. Saintekom*, vol. 11, no. 1, pp. 34–43, 2021.

[14]   M. Rizka, "Perpaduan Diffie Hellman dan Blowfish sebagai Sistem Keamanan Dokumen," *J. Infomedia Tek. Inform. Multimed. Jar.*, vol. 6, no. 2, pp. 86–90, 2021.

[15]   E. Gunadhi and A. Sudrajat, "Pengamanan Data Rekan Medis Pasien Menggunakan Kriptografi Vigenere Cipher," *J. Algoritm. Sekol. Tinggi Teknol. Garut*, vol. 13, no. 2, pp. 295–301, 2016.

[16]   N. Fahriani and I. Kurniawati, "Keamanan Data Pasien dengan Algoritma Blowfish pada HOTSPODT," *J-COSINE*, vol. 5, no. 2, pp. 140–148, 2021.

[17]   E. E. Nylis and P. Purwanto, "Implementasi Pengamanan Email Menggunakan Metode Blowfish Berbasis Web Pada UPT Puskesmas Pondok Kacang Timur," *J. SKANIKA*, vol. 1, no. 2, pp. 570–576, 2018.

[18]   H. G. Simanullang and A. P. Silalahi, "Algoritma Blowfish Untuk Meningkatkan Keamanan Database Mysql," *J. Method.*, vol. 4, no. 1, pp. 10–14, 2018.

[19]   B. H. Nuboba, I. G. N. A. C. Putra, and I. K. G. Suhartana, "Rancang Bangun Aplikasi Enkripsi dan Dekripsi Objek 3 Dimensi menggunakan Algoritma Blowfish," *J. Elektron. Ilmu Komput. Udayan*, vol. 8, no. 3, pp. 307–316, 2020.

[20]   H. Rosianto and L. Anifah, "Implementasi Algoritma DES Berbasis Blowfish untuk Enkripsi dan Dekripsi Data," *J. Tek. Elektro*, vol. 6, no. 2, pp. 121–128, 2017.

POLITEKNIK NEGERI BALI

9 772580 563008