The combination of user experience evaluation method in asessing the application of suicide risk idea identification

Tenia Wahyuningrum[1], Gita Fadila Fitriana[2], Dyah Wahyuningsih[3]
[1]Department of Informatics, Institut Teknologi Telkom Purwokerto, Banyumas, Indonesia
[2]Department of Software Engineering, Institut Teknologi Telkom Purwokerto, Banyumas, Indonesia
[3]Departement of Nurse, Politeknik Kesehatan Kemenkes, Semarang, Indonesia

Article Info

ABSTRACT

The global pandemic COVID-19 has led to extreme loss of human life worldwide. The pandemic has made a profound impact on the psyche of some people, and it has led to suicide. The need for confident handling of specific risk factors for suicide is needed. Therefore, to support Indonesian governments to control the increment of suicidal ideas for adolescents using a combination of unmoderated remote usability testing (URUT) and usability metric for user experience (UMUX)-Lite. Then we adopted the school-based mental health program to make a prototype mobile application using the risk factor for suicidal ideation (RFSI) instrument to identify suicidal ideas in adolescents. Participants' characteristics included late adolescents aged 19-22 years, male and female. The results show that the time-based efficiency on the registered task obtained from the calculation is 0.025 goals/sec. In one second, participants could complete a 0.025 of the job; although all participants could complete all tasks well and quickly, they provided good satisfaction scores. Some design improvements are needed on the prototype by considering user input.

Corresponding Author:

Tenia Wahyuningrum
Department of Informatics, Faculty of Informatics, Institut Teknologi Telkom Purwokerto
DI Panjaitan st. 128 Purwokerto, Kabupaten Banyumas, Indonesia
Email: tenia@ittelkom-pwt.ac.id

1. INTRODUCTION

The global pandemic coronavirus disease 2019 (COVID-19) has had a broad impact on human life [1], economically and educationally [2], and on mental health aspects [3]. Economic problems, uncertainty, and social isolation are associated with anxiety, stress, and suicide [4]. There has been an increase in suicides from 1% to 145% due to the COVID-19 pandemic—several surveys highlight that children and young people's mental health are disproportionately affected by adults [5]. Although suicide rarely happens in young people, it is still one of the leading causes of death in late childhood and adolescence worldwide[6]. The study of Indonesian adolescent students found that 4.75% and 2.46% reported suicide attempts or suicide attempts in the last 12 months [7]. The main risk factor for suicide among adolescents is a significant problem that must be addressed, including mental disorders, previous suicide attempts, personality characteristics, genetic loading, and family processes [6]. National and local policies are needed to create specific communications between physicians and health care delivery systems on suicide prevention activities. Prevention of suicide in the COVID-19 era requires handling particular risk factors for suicide during a pandemic and pre-pandemic [8]. Martinengo *et al.* [9], one of the efforts to improve self-prevention is to use a mobile application that can meet the unreached needs of people at risk. Another study regarding

the application of suicide ideation prevention conducted by [10] discusses the interaction of inpatient participants in filling out daily journals. Based on the weaknesses of implementing mobile applications to identify new suicide risks, we are developing a prototype that can be accessed quickly, maintain user privacy, and understand program users' needs in maintaining mental health support.

We adopted a school-based mental health program initiated by the Indonesian government as a preventive effort in adolescent health care [11], using risk factor for suicidal ideation (RSFI) instrument called SERIINA. This prototype we named suicide risk idea identification for teenager (SERIINA) [12]. The SERIINA is equipped with the involvement of the medical party as the definitive reference in handling the idea of preventing suicide in adolescents. The RFSI form items consist of four dimensions: belongingness, loneliness, hopelessness, and burdensomeness [11]. After the prototype development process is complete, the next stage is testing. In general, prototype testing will involve the user to discover the user experience. Based on previous research, testing the user experience (UX) of the SERIINA application was carried out by experts. This finding states that SERIINA has been quite good, with heuristic scores between 72-85% [12]. However, we need to listen to users' opinions before sharing the application with the market. The problem faced in testing during the pandemic is the limited space for observers and users, where there is a social distancing policy.

During a pandemic, researchers choose testing that relies on technological conducting remote research. Remote usability testing (RUT) has the advantage of budgetary savings, and it is possible to perform in a user-friendly environment. However, the downside of remote testing is reducing experimental control [13]. The unmoderated RUT (URUT) method is suitable for simple tasks with specific execution steps [14]. The URUT is a test that separates evaluators and users in different spaces and or times. Researchers do not need to schedule individual meetings with each participant in the URUT method, which takes less time than moderated studies [14]. The URUT method is the best solution to evaluate the application of SERIINA by getting fast feedback and facilitating our respondents' characteristics scattered in several parts of Indonesia. Comprehensive UX research needs to involve several quantitative and qualitative measurements. These measurements determine whether the application has met the user satisfaction standards as a whole. There are three usability quantitative and qualitative characteristics of UX that are measured; effectiveness, efficiency, and satisfaction [15]–[18]. Using URUT, effectiveness can be measured by observing task completion (task success). Sauro and Lewis [19], the average task completion rate is 78%. However, researchers can set targets more or less depending on their needs. Efficiency is measured in terms of task time: the time it takes for participants to complete the task successfully. The smaller the time required, the better the efficiency. Research by combining qualitative and quantitative methods will enrich the findings and complement each other [20], [21]. The UX score calculates by post-study questionnaire, usability metric for user experience (UMUX)-Lite. UMUX-Lite is a derivative of the UMUX questionnaire, which includes general questions about efficiency, effectiveness, and satisfaction [22], [23]. The prototype will be satisfactory if the score exceeds 76.7 [24]. The representation of the UMUX-Lite score is the same as the interpretation of the system usability scale (SUS) score. The contribution of this research is the combination of the URUT method as a complement to usability testing with the UMUX-Lite questionnaire in the case study of the SERIINA application.

2. PROPOSED METHOD

The URUT method is suitable for evaluating highly functional web applications or prototypes. The URUT methods require good planning because they cannot rely on human judgment to quickly adjust research procedures. In general, the steps for the URUT preparation consist of 3 steps: defining study goals and selecting testing software, writing tasks, and recruiting participants. In more detail, each step we describe as Figure 1.
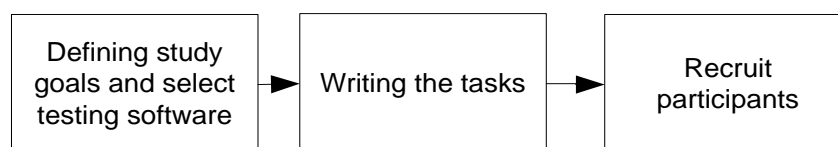

Figure 1. URUT preparation

2.1. Defining study goals and selecting testing software

In this stage, we define the study goals for the prototype and the types of participants. The study goals measure how long it takes people to complete a register, login, and complete the suicidal ideation screening (RFSI) form. The study tool must measure time on task, display individual times to inspect outliers, record screen and audio of native prototype. Researchers must ensure that the selected software is easy to access and not hinder participants' performance. For this purpose, we need a smartphone device that has a screen recorder built-in or a screen recording application. Researchers should also provide telephone or email support to participants for practice and if they encounter any problems or have questions after completing assignments.

2.2. Writing the tasks

We write task instructions for the URUT method, including what the user would like to achieve, then explain specific, realistic, and actionable instructions. Tasks should be written without including instructions that make the task too easy. This type of task instruction is adapted to qualitative or quantitative studies. Participants must achieve three main goals: i) registering an account, ii) logging in to the application, and iii) filling out an RFSI form. A task scenario is an action that asks participants to perform a series of tests on an interface. The scenario task used in this test in this test is:

"*You are experiencing discomfort. It would be best if you saw whether this feeling is normal or not. Open the SERIINA app and see if you can handle it*".

In detail, we can describe the test as; i) user goal: register an account. Task: register your name and email on the application, then fill in the required data, ii) user goal: logging in to the application. Task: enter the registered username and password, and iii) user goal: filling out an RFSI form. Task: fill out the RFSI screening form based on state feeling.

SERIINA user interface prototype using the Indonesian language can be seen in Figure 2. The registration form interface consists of 3 textboxes Figure 2(a), namely username, password, and password confirmation. If the password and password confirmation are not the same, an error message notification will appear. In the next step, after successful registration, the user is asked to write down personal data (see Figure 2(b)). The fields consist of name, age, school, and grade. Figure 2(c) shows the login interface, where the user must enter a username and password. Meanwhile, Figure 2(d) shows an RFSI form that users must fill out to screen their psychological state.
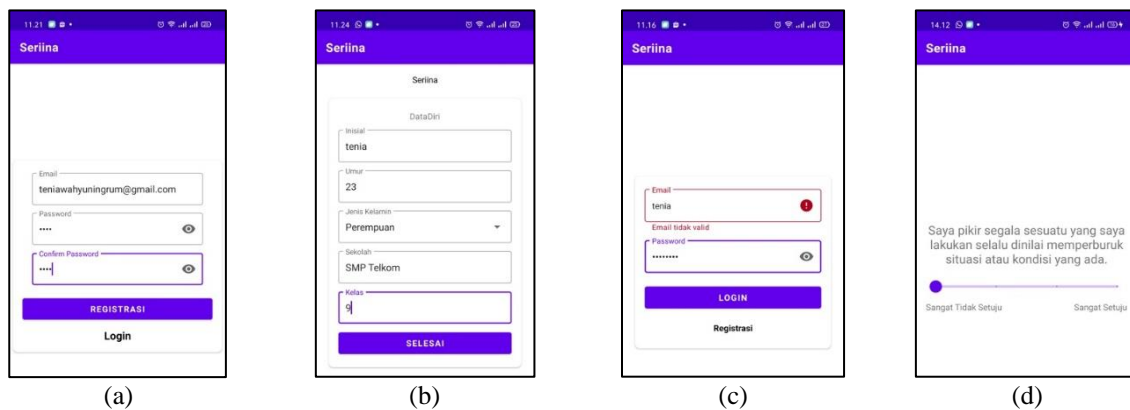

| (a) | (b) | (c) | (d) |

Figure 2. The user interface of the, (a) register, (b) user profile, (c) login, and (d) RSFI form

2.3. Recruit participants

A total of 20 participants contributed to this study. Based on Barnum *et al.* [25] opinions, five participants are not enough, mainly because they do not take individual differences into account. Participant characteristics included late adolescents aged 19-22 years, male and female. All participants have graduated from high school and are currently pursuing undergraduate studies. Participants were recruited through the Whatsapp group class on human and computer interaction courses in Institut Teknologi Telkom Purwokerto, Indonesia.

2.4. URUT and UMUX-lite combination

The proposed method in this study includes a combination of URUT and UMUX-Lite. The participants in the study will work through several stages in the unmoderated usability testing method: i) they were filling in the consent form, ii) they were doing tasks while recording activities, iii) they are filling out the UMUX-Lite questionnaire. After completing some tasks, the participant's activity recordings and questionnaires will be checked by the researchers. The researcher: i) checking form concerns, ii) they were observing video recordings, iii) Analyzing questionnaires, and iv) they were analyzing user experience. Figure 3 explains the proposed method.
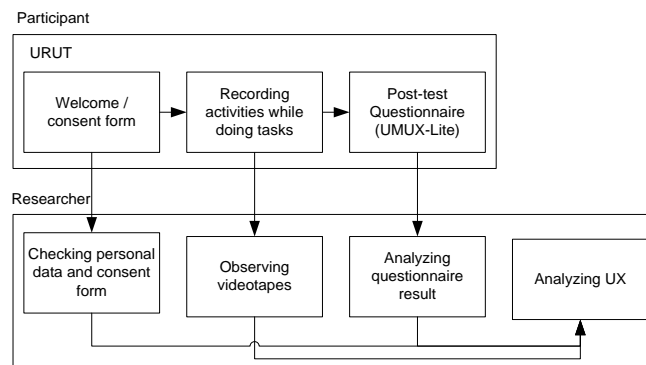
Figure 3. Proposed method

In the URUT method, participants complete the study independently, without the guidance of the researcher. Therefore, researchers need to prepare appropriate tools to collect quantitative and qualitative data. Furthermore, it is important to understand different tools can collect those different types of data. In this study, quantitative data was measured when the user was performing a task by observing the time spent on the task, the level of success, and the perceived difficulty. In addition, qualitative data such as user satisfaction was measured using the UMUX-Lite questionnaire. The URUT method is suitable for usability testing quickly and accurately to assess the exact frequency of problems; it can also be used to quickly assess participant subjective reactions from a large group of participants.

3. RESEARCH METHOD
3.1. Unmoderated remote usability testing

URUT allows the facilitator and participant interaction not to be emphasized. The researcher prepares a written assignment; the participants complete the task independently in their own time. After the participants finished the test, the researcher received the session recordings and observed the participants' performance to calculate usability metrics [14]. Figure 4 shows the URUT flow information.
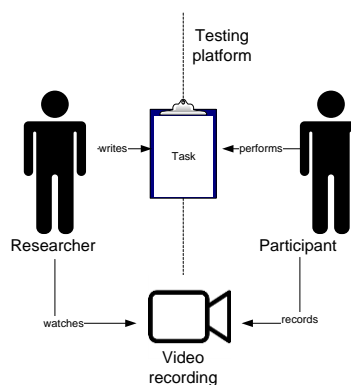
Figure 4. The URUT flow information

3.2. Risk factor for suicidal ideation

The RFSI instrument contains a fair assessment of psychological conditions, such as emotions, feelings, and adolescent judgments of their friends and environment [11]. The questionnaire consists of 16 questions. The analysis/conclusion of the screening results regarding whether adolescents are at risk of having suicidal ideation or not is a note for the competent counselor. If the total score is <34, there is no risk of having suicidal ideation (green background appear); conversely, if the score is>34, adolescents' risk having suicidal ideation (orange background appear). For more detail, Table 1 shows the initial scores and recommendations the system has given to users.

Table 1. Recommendation for youth

Variable	Recommendation
>34	i) Counseling adolescents to explore risk factors for suicidal ideation, ii) facilitate teenagers to practice stress management, and iii) They are facilitating consultation with a psychologist or health worker if needed.
<34	(i) practicing positive thinking with friends, the environment, and the situation, (ii) fill your time with valuable activities, (iii) talk about when there is a problem with someone you trust, (iv) practice problem solving, (v) use relaxation techniques when you are tense with sports, playing, recreation, worship, and (vi) don't forget to try to be happy.

3.3. Usability metric for user experience-lite

UMUX-Lite is a two-item questionnaire-based resolving some of the criticisms with UMUX. It includes only the two positive items from the UMUX but is just as reliable and valid as the UMUX [26]. UMUX-Lite is 80 percent shorter than the SUS (10 item questionnaire) and 50 percent shorter than the UMUX (4 item questionnaire). Fixes the main criticisms with the UMUX by using only the positive items from the UMUX and being a true unidimensional assessment. This reduces confusion and aligns closer to the SUS [27]. UMUX-Lite is a simplified UMUX questionnaire based on two question items related to user acceptance theory [26]. UMUX-Lite consists of two questions that imitate the technology acceptance model (TAM), measuring perceived usefulness and usability. UMUX-Lite using the 5-point scale (1 means strongly disagree, five means strongly agree), with the following two items:

Item01: The SERIINA capabilities meet my requirements (assessing perceived usefulness). (1, 2, 3, 4, 5).
Item02: The SERIINA is easy to use (assessing usability). (1, 2, 3, 4, 5).
The UMUX-Lite Score can be calculated as (1).

$$UMUX - Lite\ Score = 0.65\ (Item01 + Item02 - 2)(100/8) + 22.9 \tag{1}$$

4. RESULT AND DISCUSSION
4.1. Efficiency

Quantitative data analysis based on video observations from URUT. Observation to see user behavior and record the time. The time-based efficiency has a set of goals/seconds—efficiency computed using (2).

$$\overline{P}_t = \frac{\sum_{j=1}^{R} \sum_{i=1}^{N} \frac{n_{ij}}{t_{ij}}}{NR} \tag{2}$$

Where $\overline{P}_t$ is time-based efficiency, $\overline{E}$ is effectiveness, N is the total number of tasks, R is the number of participants, n_{ij} is the result of job i by participant j, if the participant completes the job, then n_{ij} =1 if the participant does not complete the task n_{ij} =0. t_{ij} is the time spent by participant j to complete task i.

Register task:
$$\overline{P}_t = \frac{0.507}{20} = 0.025\ goals/s$$

Login task:
$$\overline{P}_t = \frac{1.22}{20} = 0.061\ goals/s$$

Filling RFSI form task:
$$\overline{P}_t = \frac{0.830}{20} = 0.042\ goals/s$$

The time-based efficiency on the registered task obtained from the calculation is 0.025 goals/sec, or it said that in one second, participants could complete a 0.025 of the job. Time efficiency defined as how fast users work with products. The most efficient user performance on the job is the login task (0.061 goals/sec) based on the three charges.

4.2. Effectiveness

Effectiveness is calculated by dividing the number of users who complete a task by the total number of users (completion rate). Then, overall integral product effectiveness $\bar{E}$ will be calculated according to (3).

$$\bar{E} = \frac{\sum_{j=1}^{R}\sum_{i=1}^{N} n_{ij}}{NR} \times 100\% \tag{3}$$

Where $\overline{P_t}$ is time-based efficiency, $\bar{E}$ is effectiveness, N is the total number of task, R is the number of participants, n_{ij} is the result of job i by participant j, if the participant completes the job, then $n_{ij}=1$ if the participant does not complete the task $n_{ij}=0$. t_{ij} is the time spent by participant j to complete task i.

$$\overline{E} = \frac{(20x1) + (20x1) + (20x1)}{3 \; x \; 20} x \; 100\% = \; 100\%$$

Based on observation, 20 participants were able to complete all tasks successfully. So, the effectiveness can be calculated by considering the completion rate, 100% (excellent).

4.3. Satisfaction

We analyzed the UMUX-Lite questionnaire collection results based on two question items that represent perceived usefulness and usability. This study uses a Likert scale of 1-5, and the result is shown in Figure 5. The chart shows that the value of the perceived usefulness item is 3.9, and the usability value is 4.1. There was no significant difference between perceived usefulness and usability; user responses to the two question items are good, they tend to agree or be satisfied enough that the SERIINA application is easy to use. Based on Table 2, it can be seen that UMUX-Lite item numbers 1 and 2 have a correlation value of 0.946 and 0.951, which means that they are valid when compared with the value of the r table. If viewed from the significance value, with a probability of 0.05, it can be seen that the significance value is 0.000 <0.05, so it can be said that both items are valid and can be used as research measuring tools.
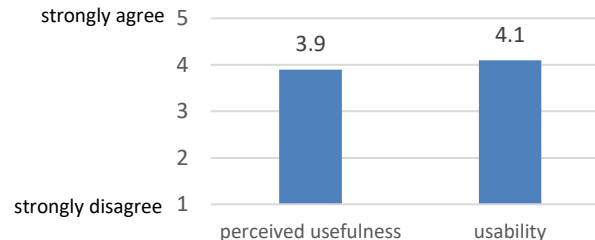
Figure 5. UMUX-lite results

Table 2. Pearson correlations

	Correlations	Q1	Q2	Total score
Q1	Pearson Correlation	1	0.799**	0.946**
	Sig. (2-tailed)		0.000	0.000
	N	20	20	20
Q2	Pearson Correlation	0.799**	1	0.951**
	Sig. (2-tailed)	0.000		.000
	N	20	20	20
Total Score	Pearson Correlation	0.946**	0.951**	1
	Sig. (2-tailed)	0.000	0.000	
	N	20	20	20

**. Correlation is significant at the 0.01 level (2-tailed).

After testing the validity, the next step is to test the reliability. Reliability testing is done to see the level of consistency of a questionnaire. The reliability test results using Cronbach's Alpha value is 0.888, from 2 question items. The basis for decision making in the reliability test, if the Cronbach's Alpha value is more than 0.6, it can be said that the question items in the UMUX-lite questionnaire are reliable and consistent. The calculation of the UMUX-Lite score is carried out using (1). The descriptive statistics of

the UMUX-Lite score are that the minimum of the UMUX-Lite score for SERIINA is 22.90, and the maximum is 87.90. Therefore, the mean value is 71.650, with a standard deviation is 0.168. Based on Sauro–Lewis curved grading scale (CGS) [24], this value equals grade C+ (Fair).

4.4. Analyzing user experience

Some things need to be improved in the design to improve the user experience, namely. User experience shows when filling out the suicide risk idea instrument, the user cannot see his position, and the system status is not visible. The solution to this problem is to add a dynamic step wizard. Step wizards will tell the user how far they have filled out the form and how many more steps will take to complete it Figure 6. In Figure 6(a) shown in initial design and Figure 6(b) shown proposed design for RFSI form using dynamic step wizard.
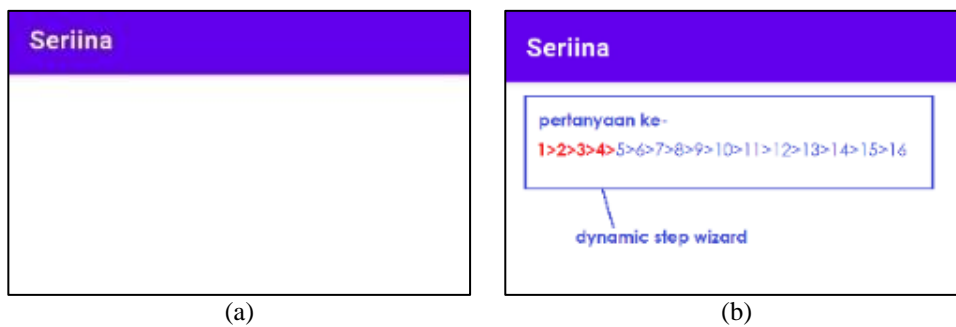
Figure 6. Analyzing user experince (a) the initial design and (b) proposed design for RFSI form using dynamic step wizard

The user also does not find the relevant icon when filling out the instrument. Icon makes it easy for users to operate the application. For example, users can access a task simply by touching the icon that represents the task. Figure 7 shows the design initiation, 7(a) explains the initial design, and Figure 7(b) illustrates the proposed design for RFSI form using the relevant icon.
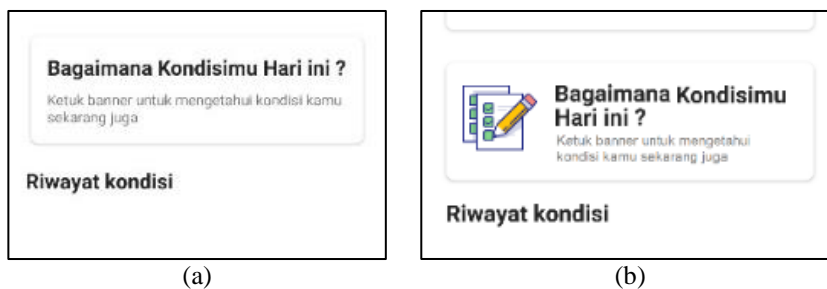
Figure 7. Design initiation (a) the initial design and (b) the proposed design for RFSI form using relevant icon

Although there is a feature to be able to communicate with counselors, but there is no feature of application use help documents, or frequently ask and questions (FAQ). User expectations are using the system or application can solve problems and work. Applications should assist users in solving their problems or tasks, so applications need to provide features and documentation that help from possible errors in use.

5. CONCLUSION

The pandemic has impacted all aspects of life, especially in the mental state of teenagers. For a study of screening of suicidal ideation in adolescents and under minimal conditions, this study involved users with an unmoderated remote usability testing method. This research combined the URUT method with the

UMUX-Lite questionnaire to add qualitative usability problem findings and optimize quantitative studies to confirm success in building prototypes. The test results show that although all participants can complete all tasks well and quickly, they provide sufficient satisfaction scores. Some design improvements are needed on the prototype by considering user input.

[25] Ni. Bevan, C. Barnum, Gi. Cockton, J. Nielsen, J. Spool, and D. Wixon, "The 'magic number 5,'" in *CHI '03 Extended Abstracts on Human Factors in Computing Systems*, 2003, no. 5, pp. 698–699, doi: 10.1145/765891.765936.

[26] J. R. Lewis, B. S. Utesch, and D. E. Maher, "UMUX-LITE - When there's no time for the SUS," in *CHI '13: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, Apr. 2013, pp. 2099–2102, doi: 10.1145/2470654.2481287.

[27] S. Borsci, P. Buckle, and S. Walne, "Is the LITE version of the usability metric for user experience (UMUX-LITE) a reliable tool to support rapid assessment of new healthcare technology?," *Applied Ergonomics*, vol. 84, pp. 1–5, Apr. 2020, doi: 10.1016/j.apergo.2019.103007.

BIOGRAPHIES OF AUTHORS

Tenia Wahyuningrum 🆔 🇬 SC Ⓟ is a lecturer in the Faculty of Informatics at Institut Teknologi Telkom Purwokerto, Indonesia. She took her undergraduate (S. Kom) at STMIK Widya Utama Purwokerto in 2005, and Master (M. T) in the Electrical Engineering Department at Institut Teknologi Bandung in 2010. He earned a Doctoral from the Department of Computer Science, Universitas Gadjah Mada Yogyakarta, Indonesia. Her research areas of interest are Human-Computer Interaction, Algorithm and programming, and Software Engineering. She can be contacted at email: tenia@ittelkom-pwt.ac.id.

Gita Fadila Fitriana 🆔 🇬 SC Ⓟ is a lecturer in the Faculty of Informatics at Institut Teknologi Telkom Purwokerto, Indonesia. She took her undergraduate (S. Kom) at Sriwijaya University in 2010 and Master (M. Kom) in the Faculty of Computer Science at Sriwijaya University in 2015. Her research areas of Soft Computing, Robotic and programming, and Software Engineering. She can be contacted at email: gita@ittelkom-pwt.ac.id.

Dyah Wahyuningsih 🆔 🇬 SC Ⓟ a lecturer in the Faculty of Nursing at Politeknik Kesehatan Kemenkes Semarang, Indonesia. She took her undergraduate (S. Kep) in 2003 and nurse profession in 2004 at Universitas Diponegoro, Indonesia, and Master of Nursing (M. Kep) in Faculty of Nursing at the University of Indonesia, Indonesia in 2009. Her research areas are Mental Health of Nursing, Community Mental Health Nursing, Rehabilitation Therapy, Occupational Therapy and Nursing Theory. She can be contacted at email: dyahwahyuningsih@poltekkes-smg.ac.id.