

DEVELOPMENT OF ARTIFICIAL INTELLIGENCE SERVICE ROBOT ACCORDING TO THE KOREAN ETHICAL SELF-CHECKLIST

JONG-WOOK KIM, YOUNG LIM CHOI, SANGHYUN JEONG,
Department of Electronic Engineering, Dong-A University
RS906 37, Nakdong-daero 550beon-gil, Saha-gu, Busan, Republic of Korea
Artificial Intelligence and Robotics Lab
E-mail: kjwook@dau.ac.kr, lotuswave@naver.com, j921sh@naver.com,
<http://deas.donga.ac.kr>

JEONG-HYE HAN
Department of Computer Education, Cheongju National University of Education
2065, Cheongnam-ro, Seowon-gu, Chungbuk, Republic of Korea
E-mail: hanjh@cje.ac.kr

This paper presents development of artificial intelligence (AI) service robots according to the Self-checklist for AI Ethics Standards announced by the Korean Ministry of Science and ICT. In this paper, we introduce the AI ethics self-checklist published in 2022 in Korea and explain development of the AI service robot to which the ethical checklist are applied. The ethical standards of the self-checklist consist of 35 items that correspond to each of the 10 core ethical requirements. In this paper, five ethical requirements applied to our home care service robot will be explained.

1. Introduction

With the development of AI after the 4th industrial revolution, AI technology has been applied to various existing services. In addition, industrial fields using AI technology were created. For example, among social media, YouTube recommends videos based on user's viewing history, and Facebook and KakaoTalk use chatbots that incorporate AI technology and are used in many business fields. Because of the popularization of AI, these technologies began to influence one's own value judgment in everyday life without realizing it. Accordingly, human beings are placed in a situation in which decision-making methods change and new values are formed.

In Korea, interest in AI started to rise in March 2016, when Sedol Lee and AlphaGo played Go [1]. Motivated by this opportunity, Korean government began to announce various policies related to AI, and in 2019, the government announced a five-year AI economic roadmap [2], suggesting that AI technology is becoming more important.

There are high concerns about the negative impact of advances in AI technology. The reason is because of gender discrimination in the Amazon AI recruitment program in 2018 [3], starting with Microsoft's racist remarks by Tay in 2016 [4] as well as the Iruda(이루다) Chabot that caused a social stir in Korea in 2020. Scatterlab's AI chatbot Iruda has been halted due to problems with personal information leakage, remarks such as demeaning the disabled, and sexual harassment in the process of building and learning data [5]. Due to these many problems, there are many voices calling for ethical consideration from the stage of AI development. Accordingly, EU and WEF (World Economic Forum) have provided various evaluation items that can be referred to in the practice of ethical principles in the area of researching AI technology and developing AI-based products or services [6,7]. In Korea, the Ministry of Science, ICT and Future Planning of Korea developed and announced the "Self-Checklist for

the Practice of AI Ethical Standards” in December 2020, through the task of “Development of AI Ethical Policy for the Realization of Human-Centered AI”.

2. AI Ethics Checklist

This AI Ethics Self-Checklist is composed of check items that can be applied universally without being specific to the major fields such as engineering, science, social science, industrial fields such as manufacturing, IT, and medical care, and the stages of planning, design and development. In addition, users are encouraged to use them according to their purpose and characteristics, and the self-checklist is not a compulsory or binding law or guideline. It respects the autonomy of companies in the relevant field and presents the checklist as a voluntary code of ethics that can respond flexibly [8].

2.1. Purpose and Recommendation

The purpose of the self-checklist was created to implement the 10 core ethical requirements for realizing the three principles of the national AI ethics standard in the process of developing and operating an AI system. The three major principles consist of the principle of human dignity, the principle of the common good of society, and the principle of purposefulness of technology. The principle of human dignity means “AI should be developed and utilized within the scope that does not harm human life as well as mental and physical health.” The social common good principle means that “The development and use of AI for the promotion of public interest should be able to improve the universal welfare of mankind from a social, national, and global perspective.”, and the principle of purposefulness of technology is “AI technology should be developed and utilized in accordance with the purpose and intention of being a tool necessary for human life, and the process must also be ethical.”

2.2. Key Requirement

As shown in Table 1, the self-checklist presents a total of 35 checklists for each of the 10 core ethical requirements: human rights guarantee, privacy protection, respect for diversity, non-infringement, public good, solidarity, data management, accountability, safety, transparency.

Table 1. Number of check lists for each ethical requirement.

Key requirements	Number of check items	Key requirements	Number of check items
Human rights guarantee	5	Solidarity	3
Privacy protection	2	Data management	3
Respect for diversity	5	Accountability	4
Non-infringement	4	Safety	3
Public good	3	Transparency	3

The use of the 10 core ethical requirements and the corresponding check items follows the voluntary guidelines reflecting the self-checklist in the process of performing tasks by members or groups who design, implement, and manufacture an AI system and maintain and manage it can be practiced in the field.

1. Human Rights Guarantee: The development and use of AI must respect the rights granted equally to all human beings and guarantee various democratic values and rights specified in international human rights law.
2. Privacy Protection: Personal privacy should be protected in the entire process of developing and utilizing AI.

3. Respect for diversity: The diversity and representation of users should be reflected in all stages of AI development and utilization, and bias and discrimination based on individual characteristics such as gender, age, disability, region, race, religion, and country should be minimized and commercialized AI should be applied fairly to all.
4. Non-infringement: AI must not be used for the purpose of inflicting direct or indirect harm to humans.
5. Public good: AI should be used not only for the pursuit of personal happiness but also for the promotion of social good and the common good of mankind.
6. Solidarity: Relations between various groups should be maintained, and AI should be used with full consideration for future generations.
7. Data management: Each data, such as personal information, must be used in accordance with its purpose, and must not be used for purposes other than the intended purpose.
8. Responsibility: Efforts should be made to minimize the damage that may occur by establishing a responsible person in the process of developing and utilizing AI.
9. Safety: Efforts should be made to prevent potential risks and ensure safety throughout the entire process of AI development and use.
10. Transparency: Efforts should be made to increase transparency and explainability to a level suitable for the situation in which AI is used by considering the trade-offs with other principles to form social trust.

3. Service Robot DORI

The care service robot DORI was designed and manufactured based on MISoRo (Moral Intelligent Social Robot) by the authors for the purpose of providing nursing care services for the elderly at home. Usually, it needs to play the role of a helper robot at home, but takes care of the elderly when there is no guardian.

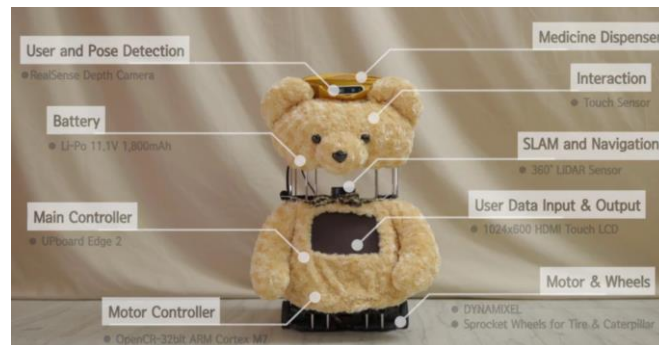


Figure 1. Appearance of the care service robot DORI

Figure 1 shows appearance of DORI, which is equipped with 2D 360° LiDAR, Intel RealSense camera, and touch LCD on Turtlebot3 Waffle Pi robot platform manufactured by ROBOTIS. LiDAR is used for SLAM and navigation required for autonomous driving, and the camera is used for user identification and face recognition, posture recognition, and object recognition required for service operation. The main controller adopts Intel NUC 11 Performance kit, and the whole system operates based on ROS, and OpenCR is used to control the wheels of the robot and manage sensory data.

DORI consists of a cognitive agent architecture Soar, a chatbot engine ChatScript, a DB management system SQLite, and open source packages for voice and image processing. Figure. 2 is a flowchart showing the process of the robot agent in understanding and processing the user's utterance. The robot basically analyzes the input information according to the 5W1H principle, and understands the user's request and the surrounding situation through the result.

When the user gives a command by voice, the voice is converted to text through Google Assistant API [9] and translated into Korean via Naver Papago NMT [10] if necessary.

If the user uses profanity or behaves inappropriately, the robot warns the user that the word or action is wrong or recommends correction of it. In addition, the robot can make its own ethical decisions based on the circumstances in which the user's human rights may be violated.

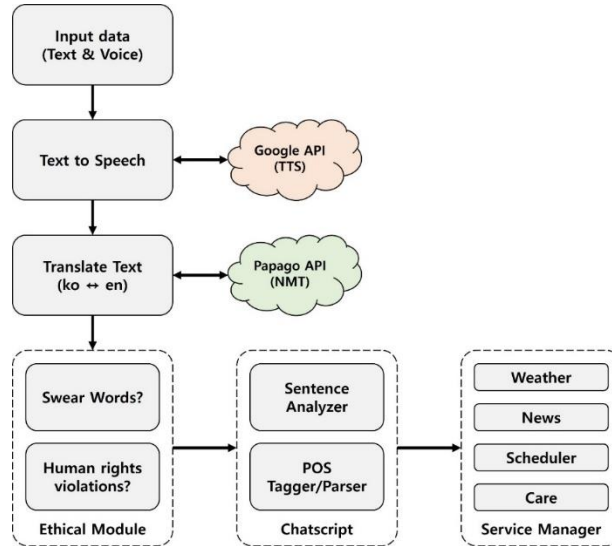


Figure 2. Flowchart of service robot DORI for speech understanding and processing

3.1. Content and Service

The service robot DORI basically has notification functions of weather, fine dust, news, and schedule management functions. As care service functions for elderly, DORI has service contents for enhancing cognitive activity, emotional activity, physical activity functions as well as medication guide and notification.

3.1.1. Cognitive activity

Cognitive function enhancement activities provide training such as guessing today's date, guessing the current year and the user's age, looking at related words and guessing names of animals and fruits, and listening to sounds and guessing transportation. The user responds to a question asked by the robot, and text cards and picture cards are displayed on the robot's screen, and the user can directly press the card to select. The robot gives the user up to three chances until they get it right, and also helps the user to find the right answer.

3.1.2. Emotional Activity

Emotional function promotion activities provide training such as recalling past memories by looking at pictures, listening to songs and expressing emotions. In this activity, the robot listen to the user's words, memorize them, and reflect them in the subsequent conversation. After showing a picture of the user's past, ask questions about that time. Then, after listening to music, robot asks about the favorite colors, places, and the people related with the music. At this time, when the user memorizes the user's words and has an activity to express emotions out loud, the user is trained to say what they like and express their emotions. Figure 3 is an example of conversation between user and robot during the emotional function enhancement activity.

```

<<Play music>>
misoro: Is there a time when you think of this song?
user: yes
misoro: Great! I wonder when it is.
user: It was when I was in elementary school
misoro: When you were in elementary school! It's interesting!
misoro: What color of clothes made you feel good?
user: yellow
misoro: I like yellow, too.
misoro: Where did you go most often at that time?
user: playground
misoro: Who did you go to the playground with?
user: Minji
misoro: You went with Minji.
misoro: Let's express our emotions together!
misoro: Clap clap clap clap 3 times and follow along
misoro: Clap clap clap! I felt good when I wore yellow
user: Clap clap clap! I felt good when I wore yellow
misoro: That makes me feel better, too.
misoro: Today, we had a time to recall and express good memories.
misoro: Shall we shout it out again?
misoro: Clap clap clap! I felt good when I wore yellow
user: Clap clap clap! I felt good when I wore yellow
misoro: Good job. I feel great, too!
misoro: A lot of good things will happen in the future.
misoro: I'll see you again next week in good health!

```

Figure 3. Example of conversation between user and robot as an emotional function enhancement activity.

3.1.3. Physical activity

In the activity to promote physical function, the game “Mugunghwa Flower is Blooming” is provided as shown in Figure 4. When the robot turns backwards, the activity starts, and the voice “Mugunghwa flower is blooming.” is played. When the voice playback is finished, DORI rotates in place to look at the user and then observes the user's movement. If no motion is detected, turn around and repeat this process. If motion is detected, the activity is terminated.



Figure 4. Example of physical activity program playing with “Mugunghwa Flower is Blooming”.

At this time, posture recognition is used as a tool to observe the user's posture and movement in body function enhancement activities. The posture recognition function was implemented using the Google AI framework MediaPipe Pose [11]. MediaPipe Pose is an ML solution for high-fidelity body pose tracking that leverages BlazePose research supporting the ML Kit Pose Detection API to infer 33 3D landmarks from RGB video frames and a background segmentation mask on the whole body. Through the learned model, the user's skeleton is recognized based on the user's face from camera data.

3.1.4. Medication guide and notification

The medication notification service provides a function to notify the medication time at an appropriate time, and provides a function to notify the guardian through a mobile messenger whether medication has been taken or not. The medication time is stored and managed in the DB for each user, and when the set time comes, the user is identified through face recognition.

If the user's face is matched, the medicine is delivered and a medication completion message is sent to the guardian. If the user refuses to take the drug, DORI persuades the user by giving a reasonable justification. However, if it refuses more than 3 times, the service is stopped after notifying the guardian.

For face recognition, a KNN (K-Nearest Neighbor) machine learning model was used [12]. A model that has learned a user's face in advance is created, and a user is specified using real-time image processing and a learning model.

3.1.5. Monitoring Service

The robotic monitoring service plays a role of helping the user to live safely in the absence of his or her family or caregiver. DORI tracks and observes users in real time as soon as the service starts. It receives the image data and depth value from the camera to determine the user's location and the distance to the user. The direction of the robot is continuously controlled to align the user in the middle of the image frame, and the distance from the user is maintained at about 1~2m based on the pelvic coordinates obtained from MediaPipe Pose and depth camera. The video observed by the user is stored in the DB, and when an emergency situation occurs or the user needs help, the message is immediately notified to the guardian so that a quick response is possible. Figure 5 shows the user's location and distance between services.

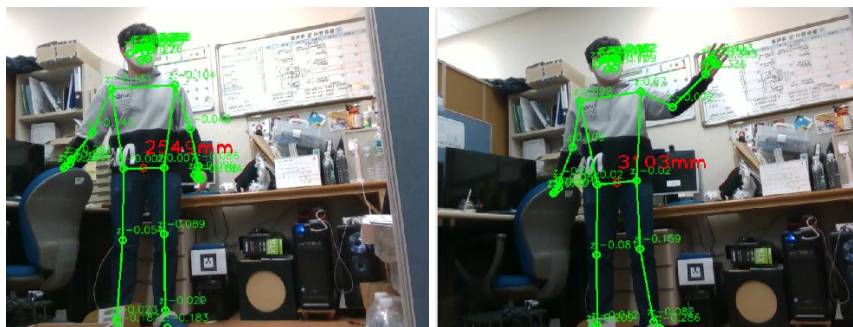


Figure 5. Measurement of the user's location and distance using MediaPipe Pose and an RGBD camera.

4. Application of AI Ethics Checklist

4.1. Cognitive and Emotional Function Promotion Activities

Cognitive and emotional function promotion activities help users to improve their emotional stability and emotional function. Robots basically use honorifics to communicate with users in a respectful way. In addition, discriminatory remarks are not allowed according to the user's individual characteristics. This reflects the first (Human Rights Guarantee) ethical requirement, and in case of Korean, respectful words are used through English-Korean translation with Naver Papago NMT.

If the user uses profanity in the course of communication, the robot informs the user that the word is a bad one that can hurt the other person and instructs the user not to repeat it again. This is due to the fact that using profanity negatively affects not only others but also the user. Frequent use of profanity intensifies emotional anxiety and increases the likelihood of violent

behavior or thinking, which in turn increases the user's aggression. This reflects the fifth (Public good) ethical requirement, and users also need to form a relationship that treats robots with respect just like humans.

4.2. *Activities to improve physical function*

In body function enhancement activities, the robot uses a camera to capture the user's movements and performs real-time posture recognition. In this case, the image data is not stored separately and is immediately removed after image processing to reflect the second ethical requirement (Privacy Protection).

4.3. *Medication Reminder Service*

In the medication notification service, the robot uses a camera to recognize the user's face for identification. The user (information subject) or guardian (third party) receives a video of the user's face to be used as learning data, and at this time, consent should be obtained from the provider according to the appropriate procedure. It also restricts data access between users by identifying users through facial recognition.

If the user ignores the medication reminder or refuses to take the medication, DORI explains the basis for medication guidance and justifiable reasons. If the user still expresses refusal, DORI says that if he or she refuses one more time, it can notify the guardian. If the user accepts taking the medicine, the medication notification service is provided as normal. However, if the user stubbornly refuses to take medicine, the current situation is communicated to the guardian and the service is stopped. This reflects the first and the fourth (Non-infringement) ethical requirements, because not performing the medication service will harm the user's health.

In addition, an accident must be prevented by providing a device or function that the user can refuse in a situation where the wrong drug is given due to an unexpected problem or dosage exceeds the daily limit. This function corresponds to the ninth ethical requirement (Safety).

4.4. *Monitoring Service*

The monitoring service tracks and observes the user in real time as soon as the service starts and informs the user to take a picture with a camera. This is a procedure for protecting the user's personal information and privacy that falls under the first (Human Rights Guarantee) and second (Privacy Protection) ethical requirements. In this case, if the user expresses a denial of service by saying "Do not follow" or "Turn off the camera", the process is immediately stopped. In consideration of the socio-economic impact corresponding to the fifth (Public good) ethical requirement, the service for elderly care robot is designed to reduce the burden on nursing and care workers and caregivers in an era of low fertility and aging.

After explaining the justification and necessity, the user's intention is checked once again. If the user says "no" again, DORI respects the user's will and stop recording the camera. It is an action based on good reason to protect the user, but should be able to abort the process if the user refuses to do so. This reflects the first ethical requirement.

Due to the characteristics of the elderly care robot, it continuously follows and observes the user, but the user may feel like being watched by the robot. To alleviate this feeling, DORI needs to maintain a distance from the user while observing for the Privacy (second) and Safety (ninth) requirements, and immediately informs the guardian of the current situation only when a dangerous situation occurs. Under normal circumstances, a third party, including a guardian, cannot know the current appearance and situation of the user.

Consequently, the authors have considered a total of five ethical requirements from the development stage of home care service robot to make a trustworthy AI robot like the following:

- 1st (Human rights guarantee): commonly crucial for robotic home care service

- 2nd (Privacy protection): removal of face and body image data
- 4th (Non-infringement): continuous request for medication
- 5th (Public good): prohibition of abusive language
- 9th (Safety): keeping distance between robot and human during monitoring

5. Conclusion

This paper presents consideration of Self-checklist for AI Ethics Standards announced by the Korean Ministry of Science and ICT in developing an elderly-care service robot. In this paper, we introduce the AI ethics self-checklist published in 2022 in Korea and explain development of the AI service robot to which the ethical checklist are applied. The authors have considered a total of five ethical requirements (first, second, fourth, fifth, and ninth ones) from the development stage of home care service robot to make a trustworthy AI robot. As future work, the authors will define detailed checklists and user experiments for these five ethical requirements which is necessary for developing and validating ethical home care service functions.

Acknowledgments

This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education(2020R1A2C1014649) and also supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. NRF-2021R1A4A1022059).

References

1. <https://www.dongascience.com/news.php?idx=10865>
2. <https://www.korea.kr/news/pressReleaseView.do?newsId=156366736>
3. <https://www.bbc.com/korean/news-45820560>
4. <https://www.yna.co.kr/view/AKR20160325010151091>
5. <https://m.lawtimes.co.kr/Content/LawFirm-NewsLetter?serial=170076>
6. The Assessment List for Trustworthy Artificial Intelligence (ALTAI), High-Level Expert Group on AI, EU, 2020
7. Companion to the Model AI Governance Framework, WEF, 2020
8. <http://www.aitimes.com/news/articleView.html?idxno=141704>
9. Google Assistant, <https://developers.google.com/assistant>
10. Papago NMT, <https://developers.naver.com/docs/nmt/reference/>
11. MediaPipe Pose, <https://google.github.io/mediapipe/solutions/pose.html>
12. Onel Harrison, Machine Learning Basics with the K-Nearest Neighbors Algorithm, <https://towardsdatascience.com/machine-learning-basics-with-the-k-nearest-neighbors-algorithm-6a6e71d01761>