REPORT

Training AI Models on Mammography Dataset for Cancer Detection, What Could Be the Ethical Issues and How to Address Them?

Haoting Chen

ANU | U7227871

Training AI Models on Mammography Dataset for Cancer Detection, What Could Be the Ethical Issues and How to Address Them?

1 Introduction

The rapid advancement of AI has been revolutionising many industries, including healthcare. A recent study conducted by Eisemann et al. shows the remarkable result that radiologists assisted by AI can pick up one extra case of breast cancer in every 1000 screenings while their workload is reduced (Eisemann, et al., 2025). The experiment was performed nationwide in Germany by comparing the results of 260,739 women whose mammography images were examined by AI plus a radiologist and 201,079 women whose images were examined by traditional double reading by two radiologists (Eisemann, et al., 2025). The specific AI models used in the experiment came from the commercially available German AI system Vara, which is based on the deep Convolutional Neural Network (CNN) structure and is trained on numerous mammography images (Eisemann, et al., 2025). Despite the achievement of AI in potentially saving a non-negligible number of human lives, the *training* and *application* of Vara AI in mammography may lead to plenty of ethical concerns, e.g., if data used to train AI comes from appropriate sources. There are two aspects to the discussion of ethical issues. One is the training of AI. The other is the use of AI. The former can also be considered a *Data Mining* process that extracts knowledge from a large mammography dataset to create an AI system equipped with the knowledge to detect breast cancer on unseen mammography images. The latter is then the outcome of the data mining process. The ethical issues aroused from both aspects are discussed. The remaining parts of the article are structured as follows. Section 2 provides a more detailed background to the data mining process for creating Vara. Section 3 discusses several key ethical issues that may occur during and after creating Vara. Section 4 provides potential approaches to address the issues, mainly from the technological perspective. Section 5 concludes the article.

2 Background

The project Vara mines mammograph image data to create an AI system that takes the mammography images of an individual as input and outputs a confidence score between 0 and 1, indicating how likely the individual has breast cancer (Eisemann, et al., 2025). The project has two main objectives: one is to automate the initial screening by operating with high sensitivity to reduce the workload of radiologists, and the other is to help radiologists pick up the positive cases they miss, especially in secondary screening (Vara, 2025; Eisemann, et al., 2025). Both were successfully demonstrated achieved by Eisemann et al.'s study (Eisemann, et al., 2025).

The nature of the data source for this project is images since the AI takes images only. However, it cannot be ignored that, during the initial data collection process, other personal information may also be collected, such as gender, age, and skin colour. The data source of this project comes from more than two million mammography images contributed by Vara's trusted partner (Eisemann, et al., 2025; Vara, 2025). Although the specific trust partners are not mentioned by either Vara or Eisemann et al. (Eisemann, et al., 2025; Vara, 2025), they are likely radiology departments around the world that provide first-hand data or institutes and researchers who have cleaned, integrated, and selected publicly available radiology data for other researchers for secondary use. This assumption is based on the common practice in AI-driven medical research, which typically collaborates with hospitals and utilises publicly available datasets. Since Vara claims their data is from trust partners, it is also assumed that most data collected to build the AI has granted formal informed consent from each individual who contributed their own mammography images.

The method of this data mining project is to train CNN classifiers to learn the features or signs in mammography images related to breast cancer so that the knowledge about how to examine mammography images to predict breast cancer is mined in stored in AI (Eisemann, et al., 2025). The CNNs also learn how to convert raw mammography images into machine-readable feature vectors from the dataset (Du, et al., 2020).

The main expected use of the mining result, i.e., Vara, is to assist radiologists in examining images to reduce their workload and false negative cases. Since Vara also takes mammography images as input during its use, the main use of Vara also involves data of a personal nature, i.e., the mammography of their body, and hence, informed consent is required. In the recent study by Eisemann et al. on the efficacy of Vara, informed consent was waived by the Ethics Committee of the University of Lübeck. Even so, participants are still informed about the use of AI at every screening site, which demonstrates good practice in aligning with the United Nations Educational, Scientific and Cultural Organization (UNESCO) principle of transparency (UNESCO, 2021) and sets an example for others who use Vara and relevant AI. Besides the main use of Vara, since Vara is based on CNNs and CNNs have also learned how to extract features from mammography images in addition to classifying these features during the data mining process, most parts of the CNNs can be used to train other relevant models, such as detecting other diseases from mammography images. Such reuse and finetuning of pre-trained models are common in machine learning and have helped accelerate the growth of machine learning and AI. In this secondary use, additional informed consent is required because the data that is originally collected for creating Vara may not be illegally used for other purpose unless the data owner makes further consent. While the data is learned by CNNs as parameters and is typically human-unreadable, it is possible to make these human-unreadable parameters more interpretable (Du, et al., 2020), potentially reflecting some person's nature. If the original input dataset is not properly anonymised, such as creating noise (Taylor, 2025). This is not a desirable situation because AI is now directly learning and each real individual. The features learned by CNN, if they can be interpreted in a human-readable way, may also help humans discover new ways to examine a mammography image that has been discovered by AI but not by humans. This secondary use may contribute to the improvement in medicine. Likewise, informed consent is required for any secondary use as the owner of data may only consent to allow their data to be used in particular ways. In addition, feature vectors and the model's parameters may also maintain some person's nature, provided the AI model is interpretable and the anonymisation of AI during training is not sufficient.

The main positive impacts of the result are the reduced workload of radiologists and the increased opportunity to save more lives since AI can help radiologists pick up some cases they miss. However, the negative impacts are evident. With the reduced workload of radiologists, employers may employ fewer radiologists, resulting in job loss. In addition, radiologists may become completely reliant on AI without proper human participation and cross-checking (Lange, 2025). Moreover, AI may constrain humans, in this case, radiologists, to become more experienced to potentially save more lives (Valenzuela, et al., 2024).

3 Ethical Concerns

Many ethical issues arose during Vara's development and use. This article discusses the most significant ones in relation to UNESCO and the Australian Privacy Principles (APP).

3.1 Safety and Security

Ensuring individual life safety is always the top priority in most scenarios. Vara AI, the outcome of the data mining project, assists radiologists in detecting breast cancer cases. However, whether a person with breast cancer can be successfully detected, especially in the early stage, causes a significant impact on their survival. Therefore, the Vara itself needs to be safe, e.g., having a prediction accuracy that meets some authoritative standard, and the operation of AI needs to be secure, e.g., preventing the parameters of the AI system from corrupting to impact the prediction accuracy of the AI. Safety and security are highlighted as UNESCO's second principle of AI. As mentioned in the study by Eisemann et al., Vara has obtained CE certification and can be used in medicine (Eisemann, et al., 2025). However, according to Vara's website (Vara, 2025), Vara is still undergoing frequent further development. Therefore, whether each updated version of Vara can be applied safely remains a concern since each development cycle may not necessarily lead to improvement but may instead introduce more bugs.

Although AI safety is important, ensuring safety by setting extremely high standards to prevent less-than-optimal AI from being used in practice may also hinder the development of AI and science. This is because practical experience is also the major source of technological development.

3.2 Privacy

As stated by UNESCO, respecting, protecting, and promoting human rights and freedom is a major value that every actor involved in AI should practice (UNESCO, 2021). Human rights include the right for every individual to protect their private data. Therefore, ensuring privacy becomes UNESCO's fifth principle of AI (UNESCO, 2021). However, from the initial data collection of Vara, the data mining process for creating Vara to the use of Vara, privacy issues may occur in all places. Even though Vara has obtained informed consent from every individual from whom data is collected, data breadth often happens unexpectedly and affects individual privacy.

During the initial data collection process, data privacy is maintained through the informed consent of individuals who supply their personal data and ensuring the quality of personal information. While the former has been briefly discussed in the Background section, the latter is also essential and is the 10th principal (APP 10) in APP (Australian Government, 2018). In the Vara scenario, this can be related to whether a mammograph image from a person with cancer is indeed labelled positive or whether the mimeograph image of each individual is not corrupted. The respect for the integrity of personal data not only contributes to the development of AI but also protects human rights and dignity since each individual and their data is treated as a whole (UNESCO, 2021). Although Vara claims their data is from trusted parties (Vara, 2025), whether or not this data is correct and of high quality may not always be checked by Vara or the data providers. However, ensuring the integrity and quality of a large amount of data may not always be feasible. Informed consent is related to the APP principle of collecting personal information by lawful and fair means (APP 3) (Australian Government, 2018), e.g., not stolen from hospitals or bought from sources of unknown legality, as well as notification of data being collected (APP 5) (Australian Government, 2018). Although Vara claims their sources are trusted (Vara, 2025), there may be edge cases, e.g., the collection of personal data can be done without consent in some countries but is strictly required in Australia and Germany, which raises issues of whether this data can be legally used by Vara, or it should follow the broader international laws.

During the cleaning, integration, and selection process, the data quality (APP 10) should be ensured, i.e., the processes should increase the quality of data instead of destroying it. However, in practice, some data cleaning and integration techniques may inevitably incur some data loss while improving the data in other aspects. Therefore, ethical concerns remain. While integrating data around the world, the potential to send personal information overseas (APP 8) need to be cautious (Australian Government, 2018). Every country may have different privacy laws. APP 8 states specifically for Australia that Australian data sent overseas should remain following all APPs during the transition or in the recipient (Australian Government, 2018). Similar regularisation may apply in other countries. After data has been pre-processed for data mining, the security of personal information (APP 11) needs to be ensured during the data storage, which is also part of the privacy concerns to prevent personal data breaches. However, since Vara is a commercial AI system and is likely closed-sourced (Vara, 2025; Eisemann, et al., 2025), how data is processed and mined beyond the initial data collection is unknown and is also confidential to the company unless it must be revealed for investigation in some situations by law. Therefore, data privacy can be a concern during the data pre-processing step. However, requiring all AI projects to be open-sourced is not feasible because AI companies need income and better sales of AI in the market allows companies to invest more in the development and advancement of AI.

During Vara's training or data mining process, the privacy principle applies in the way that entities need to use the data for which it was collected (APP 6) (Australian Government, 2018). If an individual only consents to share their data for basic statistical or medical analysis, then it may not be used to train large-scale AI models like Vara. However, since Vara's data is from various sources, it can be hard or even infeasible to ensure that all data owners have consented to their mammography images being used to train AI. During Vara's, the training team may also need to decide whether to anonymise the input data, such as creating noise (Taylor, 2025). Such action can prevent the personal information of each real individual from being learned or recorded by AI while maintaining the pattern of the whole dataset to achieve the same outcome of data mining. However, these ethical practices are not documented or provided by Vara. Therefore, this area of data privacy remains a concern.

After Vara's training or knowledge presentation process, APP 6 continues to apply, especially for the secondary use of the data collected. The Vara company may disclose information to other companies to build AI or for other purposes. However, the secondary use of personal information also requires the consent of the individuals from whom data is collected. In the worst scenario, the medical data may be illegally sold to insurance companies to set prices tailored to each individual, e.g., forcing individuals who have cancer or likely have cancer to pay a higher price, to maximise their profits. The operation of Vara is also subject to privacy measures because the input of Vara is still personal mammography images. For example, individuals who have consented to the use of AI in their mammography screening do not necessarily mean they have consented to share their data for further training and AI improvement. They just allow their data to be read only by AI. This is also related to APP 6, the use and disclosure of personal information (Australian Government, 2018).

3.3 Fairness and Non-Discrimination

Fairness and non-discrimination are the 3rd principles outlined by UNESCO. It can be understood in the way that AI technology should be available and accessible to all, regardless of gender, skin colour, region, etc. This can be further connected to UNESCO's value of diversity and inclusiveness (UNESCO, 2021). However, the Vara clearly cannot benefit every individual in the world. For example, Vara focuses on detecting breast cancer from women's mammography images (Eisemann, et al., 2025). Although it is rare for men to have breast cancer, it is not impossible. Medical advancements always need to involve rare diseases in order to benefit all of humanity. Since Vara is a commercial AI system instead of open-sourced, hospitals in low-income counties may not be able to afford the AI system to assist their examination and treatment. However, low-income and developing countries are usually the ones that urgently need AI to improve their healthcare system because there may also be a lack of human radiologists. In addition, developing countries typically lag behind in AI development due to the limit in expense and specialists. Therefore, they also insist on open-sourced AI models to catch up with AI development. Moreover, since the source of Vara comes from places where radiographic imaging equipment at least exists, the dataset does not represent individuals from all over the world, but instead, it mainly focuses on Germany or some other developed countries (Eisemann, et al., 2025). Therefore, the data mined, i.e., the AI system built, may not always work for all human beings since it can lead to domain adaptation problems in machine learning. Non-discrimination can be interpreted as preventing bias and discrimination in humans recorded in the dataset or the unbalance of the dataset from being also learned by AI. For example, people in minority groups have low cancer rate may only be because the doctors there are less experienced, not willing to treat people in minority groups carefully, or just sample collected has coincidentally low positive cases.

However, ensuring fairness can be difficult because a single piece of AI research typically solves a program in a limited domain. Community and collaboration are usually needed to ensure the balance of AI research in all domains.

3.4 Transparency and Interpretability

Transparency and interpretability is the 7th principle outlined by UNESCO. Transparency in the Vara scenario refers to the requirement that individuals need to be fully informed when their personal mammography images are used to train or be read by AI and how their data is processed by the AI to become the final output. The latter further relies on the interpretability of AI. In Eisemann et al.'s study of the clinical experiments on Vara, all participants in the AI group, i.e., whose results were processed by AI voluntarily, consented to give their results to AI. In contrast, those who did not want to share their results with AI became the control group. This was good practice for transparency, and it respected the individual's freedom to decide whether or not they want AI to participate in their mammography screening. However, the transparency of the internal working principle of Vara may not be enough. This is not only due to the close-sourced nature of Vara but also the complexity of Vara's underlining CNN architecture.

Traditional AIs, such as symbolic AI and decision trees, typically have clear and human-readable internal structures showing how they gradually process inputs into desired outputs (Du, et al., 2020). For example, symbolic AI has a well-formed algorithm guiding all its actions, and decision trees have branches that show each decision made by AI at each stage of processing the inputs (Du, et al., 2020). However, there is a trend that the more advanced AI is, the less interpretable it is because even the process that was originally considered part of the development of AI, such as handcrafting features of input data, has been automated by AI by now, such as CNN (Du, et al., 2020). Therefore, people may no longer know what features of mammography images have been extracted to contribute to analysis and output. AI is becoming more and more like a black box (Du, et al., 2020).

Although transparency and interpretability are key ethical considerations, it is also not ideal to be overly conservative by relying solely on interpretable but low-performance traditional AI models while avoiding more advanced yet less interpretable AI, like deep learning models, because it can hinder the development of AI and potentially the discovery of more surprising, complicated, but useful models.

3.5 Other Ethical Issues

There are many ethical concerns about the Vara project. For example, training Vara and foundational CNNs typically requires a large amount of computational resources and, hence, electricity. However, if the electricity used is not originally from renewable energy, there is a potential impact on the sustainability of the environment. Environment and ecosystem flourishing is one of the key values stated by UNESCO (UNESCO, 2021). When AI makes false negative cancer detection, the responsibility for the fault may not be always clear. Responsibility and Accountability is also one of the principles outlined by UNESCO. Although Vara is developed for assisting radiologists, not replacing them, the current Vara is already capable of automating the whole screening process, from input mammography images to output results. Therefore, the proportionality use of AI, which is part of the 1st principle of UNESCO (UNESCO, 2021), need to be carefully managed. The high proportionality use of AI may not only cause radiologist to lose their jobs, as discussed earlier but also reduce human oversight and make it harder to determine responsibility and accountability when AI fails. UNESCO's 6th principle is human oversight and determination (UNESCO, 2021). It highlights that there must be always humans responsible for an AI project (UNESCO, 2021). When this principle is enforced, people may become more cautious when using AI for tasks their accountable for.

4 Solutions

Although the Vara project contains ethical issues, solutions exist to address and mitigate the impact of each ethical issue. This section discusses solutions for each of the key ethical issues mentioned earlier from mainly technological perspective. However, it also introduces procedural, governance, and educational approaches that governments or individuals can take.

4.1 Ensuring Safety and Security

Ensuring safety, from the technological perspective, is typically about increasing the performance, but primarily reducing the rate of false negatives of Vara's detection. This is because a person with cancer but not successfully detected by AI likely has serious survival issues. To reduce false negatives, AI researchers should carefully analyse Vara after it is trained to find out step by step the cause, e.g., the specific layers or neurons that potentially cause the Vara to make false negative detections or what features in the input data can confuse Vara to make wrong predictions. However, such approaches require Vara to have strong interpretability, which may not be feasible unless the interpretability issue has been solved. The improvement of Vara's performance can also be done by training it on a large dataset, which can help AI experience more edge cases to become more sophisticated. Incremental learning can be used if it is supported by Vara's internal structure and has gained consent from data owners to keep training Vara while it is being used to accelerate its development. Meanwhile, data used before, during, and after Vara's training need to be secured, e.g., undertake data encryption and data checksum and backup to prevent cyber-attacks that can damage data integrity and data corruption to ensure that Vara AI is always operated as normal and as usual. However, data encryption may slow down the model, making it less competitive than human radiologists. The backup of data requires extra storage space and if such space is not properly protected, it can introduce another weakness spot for cyber-attacks.

The procedure approach involves consolidating the double-reading procedure in the common practice in mammography screening. In Eisemann et al.'s study, the original two-human double-reading is replaced by AI plus a human double-reading (Eisemann, et al., 2025). However, at least one human participation is essential, and an alternative solution is to maintain two-human double-reading while AI becomes the third examiner to increase confidence in results. For security, Vara's researchers and users should periodically perform data backup and data integrity checks. The procedure may be regulated by Vara and the end-user organisation to ensure everyone is following it.

Governments should make policies and set up standards to determine in what situations Vara and relevant medical AI systems need to be used in practice and how well the medical AI systems need to perform to be deployed in public. Governments may also issue certification to AI models to regulate what they can do and cannot do in practice. As mentioned in 3.1, high standards may not always be ideal for AI development, and exceptions may also be carefully granted to those early-stage models with high potential to make huge achievements.

Radiologists who interact with Vara need to be educated on the potential safety and security issues of Vara. Therefore, when using Vara, they can be more cautious and avoid completely relying on the AI system. All other individuals, but particularly patients, should also be told about the safety and security of Vara. In this way, they can make a more thorough decision on whether they should trust AI to process their mammography images.

4.2 Ensuring Privacy

The technological approach to address privacy, like 4.1, also involves securing data in all stages of the Vara project and is limited to not only integrity but also confidentiality. Therefore, the data warehouse used for data mining needs to have certain encryption mechanisms as well as access control that limits AI developers and other relevant people to only access a portion of the data that they need. During the training of Vara AI, as mentioned in 3.2, the mammography images to the AI can go through some anonymised techniques to remove the biometric characteristics encoded. In this way, AI does not learn from any real individual, which ensures the outcome of the data mining process is also privacy-safe. However, security and other data privacy techniques add additional complexity to the AI models and data warehouse. Therefore, the appropriate balance is required.

The main procedure approach to ensuring privacy is to obtain more formal informed consent from the data owners regarding the usage of primary or potentially secondary data and their commitment to data security. The informed consent may also include a finite time limit on the storage of personal data to prevent future unexpected and unlimited use of personal data. The informed consent document should be clear and concise, allowing data owners to read it easily and quickly, and it should not be hidden somewhere.

Governments should ensure they have policies to enforce informed consent in data collection activities. Additional policies can be made with reference to ACC principles and UNESCO's principles, such as the more detailed policies about how personal data can be shared between organisations, how they can export to overseas countries, and what level of transparency the Vara company should provide when storing and using its data. Instead of the transparency of the AI itself, which is discussed in 4.4, transparency here is more about the disclosure of activities on the collected personal data to corresponding regulatory agencies so that the Vara company can be more self-disciplined.

Individuals who are willing to share their personal mammography data or potentially have their personal shared in an unconscious way, especially those who participate in radiologic screening, need to be educated on the importance of informed consent. Some people may directly tick all the checkboxes without reading the informed consent document carefully to save time, but this can result in personal data being shared unwillingly and unconsciously.

4.3 Ensuring Fairness and Non-Discrimination

The issues of fairness in Vara can be addressed by mitigating bias in its training data. For example, future training of Vara may also involve male mammography data to help detect breast cancer in men. The data can also be collected from remote regions, low-income countries, and minority groups. However, these groups may be less likely to access radiologic equipment and even provide their data. Therefore, the government needs to formulate policies to provide financial and technical assistance to people in these areas. Fairness and non-discrimination, from the technological perspective, are also about ensuring that the data collected is unbiased, and if it is biased, appropriate data preprocessing techniques should be used. For example, the same number of data samples from people in minority groups may be overwhelmed by people in majority groups. However, people in minority groups may have unique features in their mammography images, which others do not have but potentially contribute to the detection of cancer. Therefore, the dataset needs to be balanced so that AI can also focus on minority groups. However, shifting the focus too much on minority groups can also reduce the detection performance of the majority groups. Therefore, finding an appropriate balance is required. In addition, data from minority groups may be less accurate due to their small size and may represent the whole minority group's population. Therefore, data collectors and analysers need to carefully check the quality of data and perform subsequent collection if necessary.

More generally, the procedure to ensure fairness and non-discrimination in the Vara project is to keep examining these two properties in all stages of the development of Vara. If there are data or models that do not satisfy fairness and non-discrimination, immediate correction of data or discard of data may require to prevent the further impact in the next stage.

Governments should support remote and low-income areas to access the benefits of AI like Vara. The government and its relevant departments should conduct inspections and issue certificates for medical AI that is about to be put into clinical practice. This is not only a performance inspection mentioned in 4.1, but also an inspection of the fairness and non-discrimination of AI output results.

AI researchers need to be educated on the importance and techniques of fairness and non-discrimination so that the AI they train can achieve these two points as much as possible.

4.4 Ensuring Transparency and Interpretability

Ensuring transparency and interpretability can be hard, especially for the AI system based on CNNs like Vara, with neurons containing information that is human-unreadable. However, Du et al.'s study provides three main directions (Du, et al., 2020). The first direction is to improve intrinsic interpretability, e.g., making the model more intrinsically interpretable during the creation of the model (Du, et al., 2020). This can be done by adding interpretability constraint layers to the model, e.g., semantic monotonicity constraints, to encourage the model to use fewer features while each feature used is more related to the prediction (Du, et al., 2020). This approach can simplify the model to be more easily examined and interpreted. This direction also includes using more interpretable models, e.g., decision trees to approximate less interpretable models (Du, et al., 2020).

The second direction is to ensure posthoc global interpretability, that is, after the training of the original less interpretable model, find out the general behaviours of the model on the whole dataset, e.g., the types of features the model has extracted from its inputs that contribute to the prediction (Du, et al., 2020). Examples include a reverse-engineering approach of generating and trying different input images that can maximise the activation of a single neuron. By checking the image that maximises the activation, the specific input that a neuron is looking for can be obtained, which can further derive what type of feature the neuron aims to extract from (Du, et al., 2020). The successful adoption of this approach can not only make Vara more interpretable, as researchers can know better how each neuron is responsible for extracting what features, but also allow humans to learn from AI's learning results to improve themselves, as mentioned in Section 2.

The third direction, post-hoc local interpretability, is to learn how an AI model reacts to individual predictions (Du, et al., 2020). For example, this part of the study may discover whether marking a spot in a mammography image can change the AI prediction. If so, this spot presented in the image may be important to the AI. Through this process, humans can better learn AI behaviour and how it makes predictions based on the characteristics of input images.

Like improving security, the actions done to improve transparency can also complicate the AI system or create more tasks for AI researchers without bringing substantial performance improvements. Therefore, its necessity and degree of interpretability still need to be considered.

Other non-technical approaches involve having informed consent that AI participates in processing personal data and encourages project open sources to increase the transparency of the project greatly.

Governments, when making standards and issuing certification to AI systems, should not only consider their performance and fairness, as mentioned in 4.1 and 4.3 but also require the AI to have certain interpretability. In addition, even though a close-source project does not share its code with the public, the government can still require the project owners to subject their code to the government database for backup and examination purposes by the official.

Radiologists who use Vara should be educated to some extent by the working principle of Vara, provided that Vara is transparent enough to be looked inside. This can help radiologists better utilise Vara and take better action in the event of a failure in Vara before computer experts arrive.

4.5 Addressing Other Ethical Issues

Regarding the environmental issue of AI mentioned in 3.5, governments should encourage the AI industry to shift to clean and renewable energy through the promotion of the importance of environmental sustainability and subsidising some of the energy costs to ensure a smooth transition to new energy. Governments should also legally determine who is responsible for AI, and that there must always be a natural person responsible for AI. At the educational level, governments should encourage the appropriate use of AI rather than relying on it completely, especially in the medical field related to life safety. Even if, nowadays, AI has the ability to fully automate certain examinations, such as Vara's breast radiography screening, the law should ensure that there are natural people involved in medical examinations.

Last but not least, everyone should have awareness and literacy of ethical issues, especially in the rapidly developing fields of AI and data mining. This will help us better identify ethical issues and find solutions. Awareness and literacy are the ninth principle of UNESCO (UNESCO, 2021). At the same time, ensuring ethical correctness also requires the participation of multi-stakeholder, adaptive governance, and collaboration (UNESCO, 2021).

5 Conclusion

This article has demonstrated some key ethical issues in data mining projects of training AI on mammography datasets to detect cancer. The specific example used is the Vara project. Through Vara’s example, it was found that training and use of Vara, can lead to the key ethical issues of safety and security, privacy, fairness, transparency and interpretability. While some issues may be caused by the nature of Vara, such as its close-source nature, most issues should apply to the general field of AI in mammography. Therefore, during the future training of similar models, these ethical issues should receive special attention. Although these ethical issues are highlighted, each ethical issue is complex, and solving one could potentially raise others, such as improving security, which may complicate the model, resulting in a reduction in interoperability. Therefore, the balance point needs to be set. This article also provides the technical, procedural, governance, and educational solutions to address each issue. However, due to the complexity of each issue, the ethical problem may only be mitigated and remain existing. Therefore, solving the ethical issues of AI in mammography is a long and arduous task. As mentioned in UNESCO’s 9th and 10th principles (UNESCO, 2021), each stakeholder needs to enhance awareness and literacy and participate in cooperation to maintain ethics.

Bibliography

Australian Government, 2018. *Guide to Data Analytics and the Australian Privacy Principles.* [Online]
Available at: https://www.oaic.gov.au/privacy/privacy-guidance-for-organisations-and-government-agencies/more-guidance/guide-to-data-analytics-and-the-australian-privacy-principles

Du, M., Liu, N. & Hu, X., 2020. *Techniques for Interpretable Machine Learning.* New York, Association for Computing Machinery.

Eisemann, N. et al., 2025. Nationwide real-world implementation of AI for cancer detection in population-based mammography screening. *Nature Medicine.*

Lange, C. d., 2025. AI boost to cancer detection. *New Scientist*, 11 January.p. 8.

Taylor, K., 2025. *ANU Data Mining Week 1 Study Material.* [Online]
Available at: https://wattlecourses.anu.edu.au/mod/book/view.php?id=3353440

UNESCO, 2021. *Recommendation on the Ethics of Artificial Intelligence.* France, United Nations Educational, Scientific and Cultural Organization.

Valenzuela, A. et al., 2024. How Artificial Intelligence Constrainsthe Human Experience. *Automation in Marketing and Consumption*, pp. 241-256.

Vara, 2025. *Vara: Breast Cancer Screening Platform Powered by AI.* [Online]
Available at: https://www.vara.ai/