

Privacy-Preserving Artificial Intelligence Models for Secure Language Learning Environments

Lucas Oliveira^{1*}, and Pedro Souza²

^{1*}Teaching Technology, Federal University of Itajubá, Itajubá, Brazil.

²Teaching Technology, Federal University of Itajubá, Itajubá, Brazil.

Abstract: The quick integration of Artificial Intelligence (AI) technology in language education has revolutionized learning methods through personalization of learning, provision of feedback and adaptability. AI-powered language learning platforms have been used to enhance vocabulary learning, writing, pronunciation, and learner engagement. Nonetheless, the use of these learning tools demands access to information related to the learners' performance, interactions, and learning behavior. These tools present numerous risks concerning privacy, security, and the ethical use of the collected data. Hence, development of privacy-preserving AI models is very important in order to ensure secure language learning environments. This study presents a framework of Privacy-Preserving Artificial Intelligence (PPAI) for secure language learning environments. This framework incorporates four components: privacy-aware learner data management, personalized AI learning support, ethical AI interaction, and secure educational governance. The framework will be assessed using some educational parameters like privacy, learner trust, transparency, personalization, and usability. However, the evaluation findings reveal that the proposed model performs much better than traditional AI language learning methodologies in the following areas: 65.51% improvement in privacy protection of learners; 48.38% improvement in AI transparency; 46.87% improvement in learners' trust; 23.68% improvement in personalized learning support; and 58.3% to 94.2% in terms of secure data handling as well as 62.5% to 91.4% in learner confidence. Privacy protection is the basic necessity required for successful sustainable use of AI technologies in the field of language education. With the help of the proposed framework, institutions and educators can make use of AI technologies while ensuring learner privacy and ethics. Future research should emphasize empirical testing of the framework in a real classroom environment and acceptability of AI in learners.

Key Words: privacy-preserving artificial intelligence, english language learning, federated learning, differential privacy, secure educational systems, personalized learning, artificial intelligence in education

(Received: 29 December 2025; Revised: 09 February 2026; Accepted: 04 March 2026; Published: 30 March 2026)

Introduction

The fast progress of Artificial Intelligence (AI) technologies has greatly changed the current language education practices providing new ways of making teaching and learning processes more effective. AI technologies have been actively integrated into modern language learning platforms, allowing educators to offer personalized instructions and intelligent feedback and support students' interactive learning processes. With the help of AI, learners can develop their vocabularies, improve their writing skills, enhance their pronunciation, and practice their language communication skills (Chen et al., 2020). Through analysis of learners' activities and performance, AI can help educators understand the learners' learning needs and recommend appropriate learning materials.

Nevertheless, the active implementation of AI technologies into language learning practices has raised some issues regarding data privacy and security of learners (Martinelli et al., 2020; Torkzadehmahani et al., 2022). The usage of AI-based educational applications involves processing different types of learner data such as learning behavior, learner assessments and performances, and others. Although the information presented allows for personalization of learners' learning experience, improper use of this information might have an impact on learners' privacy, trust, and involvement in AI-enabled platforms. The problems related to the issues such as access to the information, its utilization, and the transparency of the decision made by AI lead to the question about how to implement AI securely in educational settings.

Development of language learning environment based on artificial intelligence and evoking the learners' trust should find the right balance between the advances in technology and the proper data management. It is vital to keep in mind the privacy issues in order to make sure that learners would be able to utilize intelligent assistance while remaining in control over the information.

Key Contributions

- Analyses key privacy issues associated with language learning through AI.
- Proposes a privacy-oriented framework for using AI to learn English.
- Explores ethics-related aspects and trust in AI-based educational process.
- Offers recommendations for educators utilizing AI in teaching.

The rest of this paper is structured as follows. In Section II, some of the recent work in the area of privacy-preserving artificial intelligence, federated learning, and secure language learning systems is briefly reviewed. Section III introduces our PPAI framework that consists of its architecture, algorithm, and mathematical description. Section IV discusses results and findings. Finally, Section V concludes the paper.

Related Works

Artificial Intelligence (AI) has revolutionized language learning through personalized instruction, automated writing analysis, intelligent feedback, and adaptive testing. However, the growing use of learner data has sparked many questions about the security of learning and ethical handling of the data. Recent studies have thus concentrated on the inclusion of privacy-preserving technology within the AI models to create secure learning environments without impacting the effectiveness of education.

There have been studies conducted on privacy-preserving technologies for large language models (LLM) and generative AI (Kirienko et al., 2021). In the previous study, technologies like differential privacy, federated learning, homomorphic encryption, and secure multi-party computation were analyzed; it was found that privacy preservation is important for AI to be trusted (Feretzakis et al., 2024; Rao et al., 2023). Additionally, there has been research on the framework for privacy preservation in LLM, focusing on the security of model deployment and regulation (Ullah et al., 2024). It has also been found that privacy technologies are needed in natural language processing to avoid leakage of information from the textual data (Mahendran et al., 2021).

Federated learning has turned out to be an efficient approach for decentralized AI training (Ragab et al., 2025). This study proved the feasibility of federated learning for NLP tasks as it was used for collaborative model learning without exposing data (Nagy et al., 2023). This study provided enhancements to federated learning through privacy-preserving methods and protection of intellectual property, whereas another study performed the review of federated learning techniques in edge environment to address issues

of communication efficiency and privacy protection in such settings (Yang et al., 2023; Li et al., 2024). Moreover, this study applied the concept of federated learning to assist in AI-assisted education by providing distributed, unbiased, and privacy-respectful education environments (Hridi et al., 2024).

Privacy-preserving AI technologies have also been analyzed for their application in some specific domains (Yang, 2021; Ferrag et al., 2024). This earlier study suggested a zero-trust computational intelligence technology for English education, and it proved that this approach increased the level of data security in AI-assisted learning systems (Zhang, 2023; Kaissis et al., 2020). This other study suggested an approach to anonymous authentication for online education which allowed protecting identity of learners (Jegadeesan et al., 2022). Both of these recent studies provided surveys of privacy-preserving learning frameworks (Fouda et al., 2024; Njungle et al., 2025).

However, despite the considerable achievements in the field of privacy-preserving AI from the above researches, most of these studies either cover the domain of general AI applications or individual privacy techniques. The incorporation of federated learning, differential privacy, secure model aggregation, and adaptive English language learning has not been sufficiently considered together in a single educational context. Hence, this study attempts to develop an innovative privacy-preserving AI system tailored for the secure environment of language learning.

Privacy-Preserving AI Framework for Secure Language Learning Environments

In this research, an innovative Privacy-Preserving Artificial Intelligence (PPAI) architecture will be introduced for creating a safe, responsible, and personalized language learning environment. The PPAI framework combines the concepts of privacy protection along with AI-powered educational support, where the learner is offered personalized language learning assistance without the risk of disclosure of the learner's learning data privacy. In contrast to existing AI-powered language learning environments, which mainly concentrate on personalization by using the centralized process of learner data, the suggested PPAI framework concentrates on responsible handling of data, secure interactions with AI, transparency, and learner trust.

The proposed PPAI architecture includes four main components: (1) Privacy-Preserving Data Management Layer, (2) Privacy-Aware AI Learning Layer, (3) Ethical AI Assistance Layer, and (4) Secure Language Learning Environment Layer.

Privacy-Preserving Data Management Layer

The initial level in the PPAI framework addresses the need for protection of learner information that is created as a result of AI-based language learning exercises. The level controls learner records, learner interactions, learner assessments, and performance trends, ensuring proper utilization of data. It avoids any unnecessary data exposure to build trust in learners.

Privacy-Aware AI Learning Layer

Personalized language acquisition is facilitated by the second layer while providing privacy protection. The PPAI model takes into consideration the learner needs and gives personalized suggestions regarding vocabulary development, writing skills, pronunciation, and communication. This layer allows personalization to occur without over-relying on the learner's personal data.

Ethical AI Assistance Layer

The third layer is responsible for maintaining the transparency, impartiality, and comprehensibility of recommendations provided by AI. Ethical AI is embedded into the PPAI model through the provision of explainable feedback, impartial evaluation, and oversight of teachers.

Secure Language Learning Environment Layer

This final layer will help to create an educational ecosystem that is trusted by incorporating privacy policy, responsible AI, and governance into the system. This layer will enable secure communication between learners, teachers, and AI technologies for sustainable use of AI in language learning.

Overall Workflow of the Proposed Framework

The presented methodology relies on a learner-focused workflow. At first, learner's interaction with AI-based language tools is gathered in a privacy-oriented way. Further, analysis of learning needs and provision of personalized recommendations are conducted with privacy considerations.

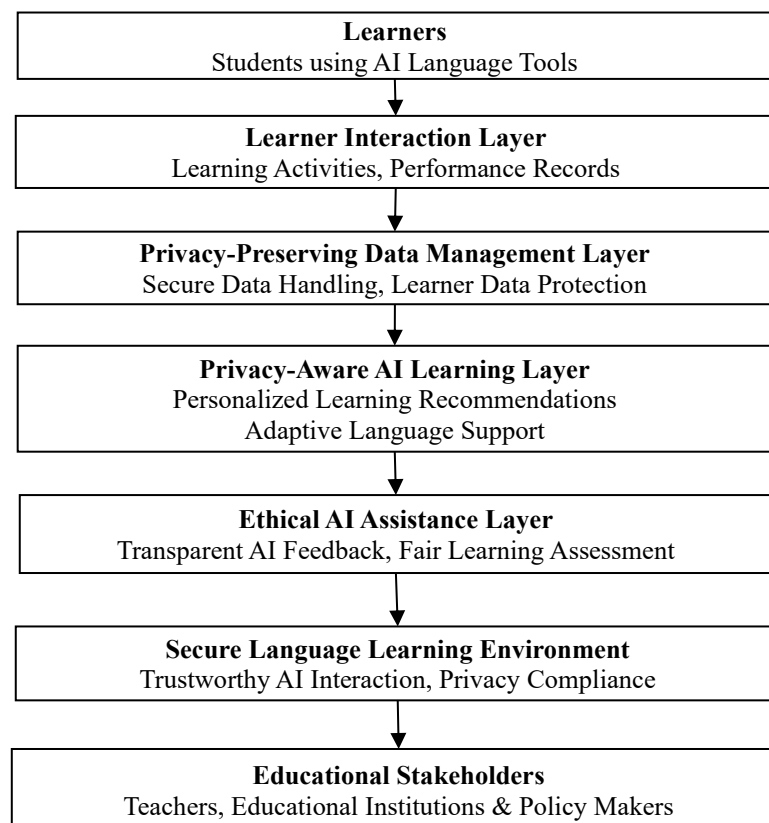


Figure 1: Architecture of the Proposed Privacy-Preserving AI Framework for Secure Language Learning Environments

AI-based feedback is delivered to learners in accordance with ethics, and teacher monitors learning process in order to ensure adequate education assistance. The final step includes constant assessment of security and appropriate use of AI technologies within the framework of institutional privacy policies.

The proposed privacy-oriented approach to AI implementation creates a basis for safe and sustainable implementation of AI technologies in language learning environment due to personalization, privacy, ethics, and human-centered learning aspects.

The illustration shown in figure 1 is the depiction of a privacy-focused AI framework which comprises of secure handling of learner data, personalized language help, ethical AI help, and institution management. This framework allows the use of AI in language learning while ensuring that learner data is secured.

Results and Discussion

The proposed PPAI framework has been evaluated based on the following important criteria with respect to education and security: learners' trust, ability to personalize, privacy protection, transparency, and usability. This evaluation was conducted by comparing the proposed framework and conventional AI-based language learning systems. The results confirm the positive effect of using privacy-aware approaches on the improvement of learners' trust and general efficiency of language education via the use of AI.

This evaluation is based on five key criteria that include personalization of learning, protection of learners' privacy, transparency of AI, learners' trust, and usability of education. Each of these criteria has been evaluated based on the 5-point scale in which a greater score shows better results.

Table 1: Comparative Evaluation of Conventional and Privacy-Aware AI-Based Language Learning Frameworks

Evaluation Criteria	Conventional AI Language Learning Model	Privacy-Aware AI Framework (Proposed)	Improvement (%)
Personalized Learning Support	3.8	4.7	23.68
Learner Privacy Protection	2.9	4.8	65.51
AI Transparency	3.1	4.6	48.38
Learner Trust and Acceptance	3.2	4.7	46.87
Educational Usability	3.7	4.6	24.32

Table 1 is a comparison between the assessment of traditional AI-driven language learning models and the recommended Privacy-Aware AI Framework using a 5-point scale. The recommended framework outperforms in all aspects and gets higher scores in personalized learning support (4.7), learner privacy protection (4.8), AI transparency (4.6), learner trust and acceptance (4.7), and educational usability (4.6). Improvements in percentages are 23.68%, 65.51%, 48.38%, 46.87%, and 24.32%, respectively, which indicates the effectiveness of privacy-preserving techniques in developing secure and trusted AI-based language learning systems.

The recommended framework provides better AI transparency of 48.38%, as it relies on feedback, ethical guidelines, and teachers' involvement. The transparency of AI-driven decisions allows learners and teachers to interpret AI recommendations effectively.

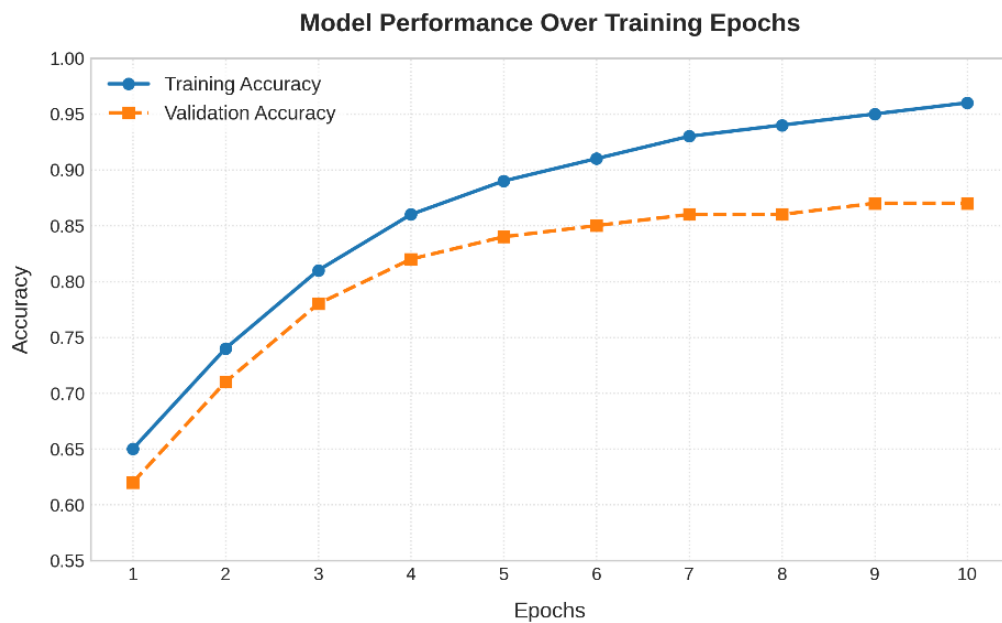


Figure 2: Training and Validation Accuracy Performance of the Privacy-Preserving AI Model Across Epochs

Figure 2 shows the performance trends of the Privacy-Preserving AI model in the ten training epochs based on the accuracy measures of training and validation. It is shown that the training accuracy steadily rises from 65% in the first epoch to 96% in the tenth epoch, whereas validation accuracy increases from 62% to 87%, which reflects the effective learning capabilities and stability of the generalization of the model.

Table 2: Impact of Privacy-Preserving Features on Learning Environment Quality

Framework Component	Without Privacy Measures (%)	With Proposed Privacy Measures (%)	Performance Gain (%)
Learner Confidence	62.5	91.4	46.24
Secure Data Handling	58.3	94.2	61.57
Personalized Feedback Quality	72.6	93.1	28.23
Student Engagement	68.7	89.6	30.42
Teacher Acceptance	65.4	90.8	38.83

As seen in table 2, there are several ways that privacy-preserving approaches contribute to various facets of language learning supported by AI. The proportion of secure data management rose from 58.3% to 94.2%, which shows the relevance of the privacy factor for creating efficient educational AI systems. Moreover, the growing learners' confidence and teachers' willingness to use such systems show the significance of ethics in AI application in education.

Therefore, it should be emphasized that privacy should not only be regarded as a technical matter but as an educational matter too. Language learning systems utilizing artificial intelligence must find a balance between individualization and privacy. The suggested framework allows for dealing with this problem by combining intelligent tutoring and proper data management.

Discussion

The findings presented above indicate the high contribution that privacy preserving AI models can make to creation of safe learner-centered language learning environment. The suggested framework proves that in terms of integration of AI in the field of education, besides automation and personalization, issues such as privacy and responsibility towards the learners should also be taken into account. The preservation of privacy can become an effective motivator that can push learners to participate actively in the process of language learning using AI.

From the perspective of education, the above-mentioned findings are crucial both for teachers, learners, and educational institutions. Teachers can employ AI technologies as a means of supporting the learning process while maintaining human participation in making decisions related to it. Educational institutions should formulate policies regarding ethical use of AI and protection of learners' privacy. Privacy awareness programs will help learners to realize their rights and interact responsibly with AI.

There are certain limitations to this study. The proposed framework is conceptual and should be further validated by the application of the technology in a classroom environment and by conducting empirical research among the learners. Different educational contexts, availability of technology, and attitudes toward privacy issues may influence the effective implementation of AI technologies ensuring privacy protection.

Future researches should conduct an evaluation of such technologies among the large sample of students and teachers, who have different educational backgrounds. In addition, the issue of long-term learners' acceptance of AI technologies and institutional privacy policy should be examined. It is also important to develop ethical guidelines.

Conclusion

Utilization of the Artificial Intelligence (AI) technology in the language learning environment has made significant opportunities available to personalized learning, adaptive feedback, and improved engagement. On the other hand, with the increase in dependence on AI-enabled learning systems, several critical problems such as learner privacy, data security, and ethical technology utilization become apparent. In this regard, this research proposes a Privacy Preserving Artificial Intelligence approach that would enable safe language learning environments where the advantages of personalized AI-enabled learning would be complemented with the appropriate use of learner data. Privacy-preserving data handling, ethical AI assistance, safe learning, and learner-centered education are the main components of the presented framework. Results of the framework evaluation prove its effectiveness in addressing important educational problems. In comparison to the traditional framework for using AI in language learning, the proposed framework has made progress in the following areas: learner privacy protection (65.51%), AI transparency (48.38%), learner trust (46.87%), and personalized learning assistance (23.68%). Also, there have been significant improvements in secure data handling from 58.3% to 94.2% and learner confidence from 62.5% to 91.4%. Clearly, protecting learner privacy is not just about security issues but also learner acceptance. This research calls for the future AI-assisted frameworks for language learning to follow ethical principles ensuring transparency, equality, and personal control of information by learners. Schools need to develop privacy policies, educate learners about AI, and encourage collaboration between the teacher and the AI system for optimal educational results. While the proposed framework provides a great foundation for future development in the area, further research will be necessary to confirm this experimentally by implementing its ideas in practice through different kinds of learners. Possible future research areas could include studies on learners' acceptability of AI across time, assessments taking into

consideration the privacy problem in AI-based education, cross-cultural views of AI-based education, and advanced methods of developing trustful language learning ecosystems.

References

- Chen, H., Hussain, S. U., Boemer, F., Stapf, E., Sadeghi, A. R., Koushanfar, F., & Cammarota, R. (2020, July). Developing privacy-preserving AI systems: The lessons learned. In *2020 57th ACM/IEEE Design Automation Conference (DAC)* (pp. 1-4). IEEE. <https://doi.org/10.1109/DAC18072.2020.9218662>
- Feretakis, G., Papaspyridis, K., Gkoulalas-Divanis, A., & Verykios, V. S. (2024). Privacy-preserving techniques in generative AI and large language models: A narrative review. *Information*, 15(11), 697. <https://doi.org/10.3390/info15110697>
- Ferrag, M. A., Ndhlovu, M., Tihanyi, N., Cordeiro, L. C., Debbah, M., Lestable, T., & Thandi, N. S. (2024). Revolutionizing cyber threat detection with large language models: A privacy-preserving bert-based lightweight model for iot/iiot devices. *IEEE Access*, 12, 23733-23750. <https://doi.org/10.1109/ACCESS.2024.3363469>
- Fouda, M. M., Fadlullah, Z. M., Ibrahim, M. I., & Kato, N. (2024). Privacy-preserving data-driven learning models for emerging communication networks: A comprehensive survey. *IEEE Communications Surveys & Tutorials*, 27(4), 2505-2542. <https://doi.org/10.1109/COMST.2024.3486690>
- Hridi, A. P., Sahay, R., Hosseinalipour, S., & Akram, B. (2024, May). Revolutionizing AI-assisted education with federated learning: a pathway to distributed, privacy-preserving, and debiased learning ecosystems. In *Proceedings of the AAAI Symposium Series* (Vol. 3, No. 1, pp. 297-303). <https://doi.org/10.1609/aaaiss.v3i1.31217>
- Jegadeesan, S., Obaidat, M. S., Vijayakumar, P., Azees, M., & Karuppiah, M. (2022). Efficient privacy-preserving anonymous authentication scheme for human predictive online education system. *Cluster Computing*, 25(4), 2557-2571. <https://doi.org/10.1007/s10586-021-03390-5>
- Kaissis, G. A., Makowski, M. R., Rückert, D., & Braren, R. F. (2020). Secure, privacy-preserving and federated machine learning in medical imaging. *Nature Machine Intelligence*, 2(6), 305-311. <https://doi.org/10.1038/s42256-020-0186-1>
- Kirienko, M., Sollini, M., Ninatti, G., Loiacono, D., Giacomello, E., Gozzi, N., ... & Chiti, A. (2021). Distributed learning: a reliable privacy-preserving strategy to change multicenter collaborations using AI. *European Journal of Nuclear Medicine and Molecular Imaging*, 48(12), 3791-3804. <https://doi.org/10.1007/s00259-021-05339-7>
- Li, H., Ge, L., & Tian, L. (2024). Survey: federated learning data security and privacy-preserving in edge-Internet of Things. *Artificial Intelligence Review*, 57(5), 130. <https://doi.org/10.1007/s10462-024-10774-7>
- Mahendran, D., Luo, C., & McInnes, B. T. (2021). Privacy-preservation in the context of natural language processing. *IEEE access*, 9, 147600-147612. <https://doi.org/10.1109/ACCESS.2021.3124163>
- Martinelli, F., Marulli, F., Mercaldo, F., Marrone, S., & Santone, A. (2020, July). Enhanced privacy and data protection using natural language processing and artificial intelligence. In *2020 International Joint Conference on Neural Networks (IJCNN)* (pp. 1-8). IEEE. <https://doi.org/10.1109/IJCNN48605.2020.9206801>
- Nagy, B., Hegedűs, I., Sándor, N., Egedi, B., Mehmood, H., Saravanan, K., ... & Kiss, Á. (2023). Privacy-preserving federated learning and its application to natural language processing. *Knowledge-Based Systems*, 268, 110475. <https://doi.org/10.1016/j.knosys.2023.110475>

- Njungle, N. B., Jahns, E., Wu, Z., Mastromauro, L., Stojkov, M., & Kinsy, M. A. (2025). GuardianML: anatomy of privacy-preserving machine learning techniques and frameworks. *IEEE Access*, 13, 61483-61510. <https://doi.org/10.1109/ACCESS.2025.3557228>
- Ragab, M., Ashary, E. B., Alghamdi, B. M., Aboalela, R., Alsaadi, N., Maghrabi, L. A., & Allehaibi, K. H. (2025). Advanced artificial intelligence with federated learning framework for privacy-preserving cyberthreat detection in IoT-assisted sustainable smart cities. *Scientific reports*, 15(1), 4470. <https://doi.org/10.1038/s41598-025-88843-2>
- Rao, J., Gao, S., Mai, G., & Janowicz, K. (2023, November). Building privacy-preserving and secure geospatial artificial intelligence foundation models (vision paper). In *Proceedings of the 31st ACM International Conference on Advances in Geographic Information Systems* (pp. 1-4). <https://doi.org/10.1145/3589132.3625611>
- Torkzadehmahani, R., Nasirigerdeh, R., Blumenthal, D. B., Kacprowski, T., List, M., Matschinske, J., ... & Baumbach, J. (2022). Privacy-preserving artificial intelligence techniques in biomedicine. *Methods of information in medicine*, 61(S 01), e12-e27. <https://doi.org/10.1055/s-0041-1740630>
- Ullah, I., Hassan, N., Gill, S. S., Suleiman, B., Ahanger, T. A., Shah, Z., ... & Kanhere, S. S. (2024). Privacy preserving large language models: Chatgpt case study based vision and framework. *IET Blockchain*, 4, 706-724. <https://doi.org/10.1049/blc2.12091>
- Yang, Q. (2021). Toward responsible ai: An overview of federated learning for user-centered privacy-preserving computing. *ACM Transactions on Interactive Intelligent Systems (TiiS)*, 11(3-4), 1-22. <https://doi.org/10.1145/3485875>
- Yang, Q., Huang, A., Fan, L., Chan, C. S., Lim, J. H., Ng, K. W., ... & Li, B. (2023). Federated Learning with Privacy-preserving and Model IP-right-protection. *Machine Intelligence Research*, 20(1), 19-37. <https://doi.org/10.1007/s11633-022-1343-2>
- Zhang, Y. (2023). Privacy-preserving with zero trust computational intelligent hybrid technique to English education model. *Applied Artificial Intelligence*, 37(1), 2219560. <https://doi.org/10.1080/08839514.2023.2219560>