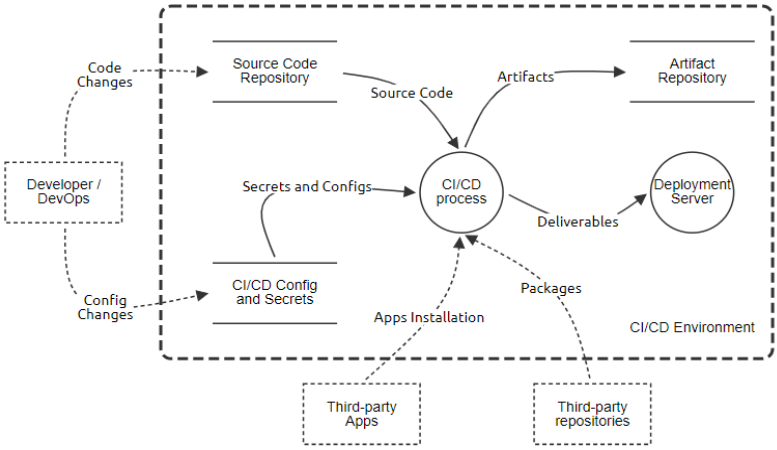

Рис.2. Структура CI/CD процесу в DFD нотації

Запропонована комбінація SSDF та STRIDE методологій забезпечує систематичне виявлення та пріоритизацію загроз в CI/CD процесах.

1. Prates, L., & Pereira, R. (2025). DevSecOps practices and tools. International Journal of Information Security, 24 (1), 11. https://doi.org/10.1007/s10207-024-00914-z
2. Paule, C., Düllmann, T. F., & Van Hoorn, A. (2019, March). Vulnerabilities in Continuous Delivery Pipelines? A Case Study. In ICSA Companion (pp. 102-108). https://doi.org/10.1109/ICSA-C.2019.00026
3. Akbar, M. A., Smolander, K., Mahmood, S., & Alsanad, A. (2022). Toward successful DevSecOps in software development organizations: A decision-making framework. Information and Software Technology, 147, 106894 https://doi.org/10.1016/j.infsof.2022.106894
4. Zhao, X., Clear, T., & Lal, R. (2024). Identifying the primary dimensions of DevSecOps: A multi-vocal literature review. Journal of Systems and Software, 214, 112063 https://doi.org/10.1016/j.jss.2024.112063

Issues of protecting personal data in artificial intelligence systems in education

UDK 37.01:004.8:004.056.5

Zhazira Yerimbetova

Kazakh National pedagogical university named after Abai, Almaty, Kazakhstan
zhazira.erimbetova@gmail.com

Currently, the digital transformation process is characterized by the active introduction of artificial intelligence (AI) technologies at all levels of the education sector. Adaptive learning platforms and intelligent systems are enhancing the

efficiency of the learning process and enabling each learner to create a personalized trajectory. However, this technological progress has also brought significant risks, the most important of which is the issue of personal data privacy and security.

Since the data is of a confidential nature, any security threat in the education sector has a high possibility of affecting an individual's future career or reputation in society. The biggest challenge of the century is building an overall data protection model for AI systems.

In the context of digitizing the educational process, artificial intelligence technologies make it possible to collect data about learners on an unprecedented scale. As noted by domestic scholar E.Y. Bidaybekov, informatization of education is not just the introduction of technical tools; it is a complex process that requires adherence to safety and ethical norms in shaping the learner's individual trajectory [1].

Foreign expert R. Luckin indicates that the possibility of collecting "invisible" data (for example, pauses in task completion, emotional responses) by artificial intelligence carries a risk of infringement of privacy [2]. This leads us to the problem of storing data in an anonymous manner. At the same time, as demonstrated in the research by C. Dwork, it has been found that "simple" anonymization does not hold up under modern de-anonymization techniques, and it is necessary to resort to "differential privacy" technology. [3].

In the space of law in Kazakhstan, personal data protection issues have been thoroughly explored in the works of R.E. Zhatkanbayeva. The author suggests that legislative regulation should be aligned with international regulation, especially regarding the provisions of the GDPR. [4,5,6].

From a technical security point of view, local researchers B.B. Akmetov and A.A. Biyakaeva suggest using cryptographic methods in the field of data integrity. This solution is in line with the idea of “Federated learning,” which was introduced by B. McMahan. This idea guarantees that the data doesn’t leave its source during the training of an AI model [7,8].

For the security of the personal data used in AI, a protection model that is comprehensive, suitable for different stages of data processing, has been developed, as shown in Table 1. The model is a combination of technical regulations and laws.

Table 1

Data security model in AI systems

Stage	Security Measure	Technology Used
Collection	Principle of Minimalism	SSL/TLS encryption
Storage	Data Anonymization	AES-256 standard
Processing (AI)	Local data processing	Federated Learning
Accessibility	Role-based access control model	RBAC system

The table takes into account the architectural features of educational platforms, and each level performs the following functional tasks:

Data collection and transmission stage. The major risk in this phase is the Man-in-the-Middle attack. By using the SSL/TLS encryption protocols, the communication channel between the learner's device and the server is secured, such that no third party can intercept the information being sent. In addition, in line with the “principle of

minimalism," only the data required for the educational process is collected, thus addressing the aforementioned risks proactively.

Data storage and anonymization. The information stored on the server should be encrypted using AES-256 (Advanced Encryption Standard). AES is one of the best encryption standards in the world. However, the information stored in the database should not only be encrypted but also de-identified to avoid the identification of the students in the future. This can be achieved by replacing the names with ID numbers. In this way, the students can be protected even if the database is compromised.

Federated Learning technology for training AI models. In traditional AI technologies, all the data is collected and stored in one place. This is a big problem. The proposed Federated Learning technology is an extremely innovative approach for education. The idea of this technology is as follows: the AI model is not trained on a central server, but rather on each individual's device (smartphone or laptop). Only new mathematical parameters are transferred to the server, and personal information is not transferred from the device. This is the highest level of privacy.

Access Management (RBAC). In an educational institution, access to data must be strictly controlled. In the Role-based Access Control (RBAC) model, for example, a teacher can only access his or her own group's progress, and an administrator can only access technical settings. This is based on the principle of "need to know."

As shown by the analysis conducted, the key problem of AI-based education systems lies in the "gap" between technological efficiency and ethical responsibility. The educational institutions should follow the "Privacy by Design" principle. This means that the security requirement should become the first need when creating any kind of software product.

The strategy for the integration of AI technology in the education sector should be accompanied by the development of digital literacy. An analysis of the research by domestic and international scholars suggests that the development of a safe digital learning environment can only be achieved by the simultaneous implementation of technical protection and legislative regulation.

1. Bidaibekov E.Y. Theory and Methodology of Informatization of Education. — Almaty, 2014.
2. Luckin, R. Machine Learning and Human Intelligence. — UCL Press, 2018.
3. Dwork, C. Differential Privacy: A Survey of Results. — 2008.
4. On Personal Data and Its Protection: Law of the Republic of Kazakhstan No. 94-V of May 21, 2013.
5. General Data Protection Regulation (GDPR). Regulation (EU) 2016/679.
6. Zhatkanbayeva, R.E. Digital Law: A Textbook. — Almaty, 2021.
7. Ahmetov B.B., Biyakaeva A.A. Information security issues in artificial intelligence systems // Proceedings of the National Academy of Sciences of the Republic of Kazakhstan. — 2022. — No. 3.
8. McMahan, B. Communication-Efficient Learning of Deep Networks. — AISTATS, 2017.