Ethical Educational Data Processing Differences of Students with Special Needs in Post-Soviet Countries

Ayaz Karimov
Faculty of Information Technology
University of Jyväskylä
akarimov@jyu.fi

Mirka Saarela
Faculty of Information Technology
University of Jyväskylä
mirka.saarela@jyu.fi

Tommi Kärkkäinen
Faculty of Information Technology
University of Jyväskylä
tommi.karkkainen@jyu.fi

ABSTRACT

Countries adopt different strategies and policies to ensure the implementation of ethical data processing and within this framework, they may encounter challenges. In this paper, we analyze the main issues regarding ethical data governance of students with special needs in post-Soviet countries and compare them. The reason why we selected post-Soviet countries as our case study region is that these countries went through different education and data privacy-related policies after the collapse of the Soviet Union. The recent Programme for International Student Assessment (PISA) results indicate that these differences can be tremendous. For example, Estonia positioned itself as the best of all non-Asian countries and economies participating in PISA, and in seventh place in the overall ranking. However, students from countries like Uzbekistan, Georgia, and Azerbaijan performed lower than average PISA scores of the Organisation for Economic Co-operation and Development (OECD) countries. This disparity highlights the important role ethical data governance plays, not only in education but also in broader socio-economic outcomes. It emphasizes the necessity of understanding these countries' various ethical educational data processing practices. Our findings show that the main challenge is bias in the data collection, where certain groups of learners are not involved. The second main challenge is about the accountability and transparency of data where most of the collected data by different organizations do not align.

Keywords

data fairness, data ethics, educational data, post-soviet countries, gem report

1. INTRODUCTION

While countries utilize educational technologies to improve the learning process, they implement various educational data governance strategies [36, 24, 10]. However, not all these implementations are successful [4]. Notably, in some

countries, issues regarding data processing, such as non-transparency and unfairness, emerge—and several of the post-Soviet countries are examples of this [15]. Following the dissolution of the Soviet Union in 1991, most of the countries went through significant changes in their education systems and applied alternative methodologies that were adopted from the Western countries [28, 2, 30]. However, after the collapse of the Soviet Union, not all member countries experienced the same level of progress [7]. While more democratic nations like Estonia or Latvia indicated the ability to provide inclusive data and publicly share their outcomes, in many other post-Soviet countries situated in the Caucasus or Central Asia, challenges have arisen in processing inclusive educational data. These differences are one of the reasons why we selected post-Soviet countries to review as their challenges in ethical educational data processing.

Differences between various post-Soviet countries can also be observed in the involvement of these countries in international data collection activities. The Programme for International Student Assessment (PISA) is one of them. While most of the post-Soviet countries take part in this assessment test, countries like Turkmenistan and Armenia did not agree on taking these tests [23]. Moreover, even some of the participating countries do not provide data that represents the average student in this country. For example, in the last edition of the PISA assessment in 2022, the tests in Azerbaijan took place only in Baku, which is the capital city [1].

Another reason to investigate the post-Soviet countries was due to the limited number of academic publications in the field of education [12]. Hernández-Torrano et al. (2021) mentioned in their bibliometric analysis that the gap in education research in post-Soviet countries is enormous. Even though the number of publications increased in recent years, the main collaborations still happen at the national level (e.g. [13, 26, 14, 8]). Moreover, they demonstrate that previous literature did not focus on the research in these regions (ibid.).

One of the goals of the United Nations (UN) is to provide equal and quality education by removing all barriers to all students in the school by 2030 [9]. Identifying students with special needs makes these individuals visible and it facilitates both government and community involvement to be able to provide the support that these learners need [25]. Understanding what types of ethical data processing practices exist can provide valuable insights for education stake-

Table 1: Grouping of Post-Soviet Countries Based on Data Governance Status

Group	Country	Description
Group 1	Estonia, Latvia, Lithuania	EU countries following GDPR requirements. These countries share data and have more complex data reporting systems.
Group 2	Moldova, Ukraine	Countries in transition to implement GDPR requirements, but not as complex as Group 1.
Group 3	Armenia, Azerbaijan Kazakhstan, Kyrgyzstan, Tajikistan, Turkmenistan, Uzbekistan	Countries following strict rules and not sharing data, making it challenging to find detailed information on data collection.
Group 4	Belarus, Russia	Countries with their regulations, similar to GDPR, providing restricted access to collected data.

holders. Thus, our concentration in this research was the data processing practices of students with special needs.

2. METHODOLOGY

We collected data from different sources. Firstly, we analyzed the Global Education Monitoring (GEM) Report, which aims to analyze the progress in the education field for reaching the United Nations' Sustainable Development Goals [34]. Since 2002, 15 various reports have been published, and their report in 2021 focused on inclusive education, where they also discuss the challenges of processing educational data of students with special needs. Thus, we selected this edition of the report to be able to review countries' policies and current strategies on data governance of learners with special needs. Moreover, we also reviewed the education policies of these countries that were shared on their official government website or international organizations' reports.

To analyze textual data from these sources, we implemented the thematic content analysis for the identification of recurring patterns [6]. Within the analysis, the codes emerged organically from the data, and we categorized the challenges related to inclusive data collection. The codes were accuracy, bias, fragmentation, and transparency whose their definitions were defined as follows [18]:

- Accuracy: correctness and precision of information collected;

- Bias: unequal representation or treatment of certain groups;

- Fragmentation: presence of disjointed or disconnected data sets;

- Transparency: accessibility of data.

Lastly, since post-Soviet countries followed different pathways in their economic growth and policy developments, their data governance strategies varied. Instead of reporting each post-Soviet country's challenges, we grouped them in Table 1 based on their status of data governance. All of the Group 1 countries are part of the European Union, thus, they follow the General Data Protection Regulations (GDPR) requirements. These countries share their data and have more complex data reporting systems. Group 2 countries are in the transition to implement GDPR requirements; however, they are still not as complex as Group 1 countries. Group 3 owns its regulations, which are similar to the GDPR and provide restricted access to the collected data. Group 4 countries follow strict rules, and they do not share the data. As a result, it is challenging to find detailed information on data collection in countries belonging to Group 4.

3. RESULTS

The main challenges in ethical data processing in most of the post-Soviet countries are transparency, bias, and accuracy (Table 2). Based on the analysis of official reports from 15 post-Soviet countries, 12 countries face accuracy-related challenges, 13 countries report bias-related issues, transparency problems are noted in all 15 countries, and fragmentation challenges are observed in 9 countries.

Group 1 countries reported that both Estonia and Latvia provide transparent and accountable data processing where Estonian schools register educational data of students with special needs and this process is regularly monitored by the External Evaluation Department of the Ministry of Education and Research [17]. In case of any issues, this department advises the schools. In Latvia, the main challenge is the detailed data of learners with special needs for tracking their academic performance [35]. UN Committee report of Latvia October 2017 also highlights the importance of possessing disaggregated data on learners with disabilities [22]. GEM Report also mentions that data quality should be improved for their usage in the policy development. In the case of Lithuania, schools do not provide accurate data on non-attendance, and in the GEM report 2021, the main challenges that they mentioned were mainly the involvement of more special education experts in rural areas and better communication with them where more complex data collection would be needed [5].

Both Group 2 countries, which are Ukraine and Moldova, reported that the main challenge is the cooperation and collaboration between government units because they do not have

a unified database that demonstrates data on learners with special needs. GEM Report for Moldova particularly mentions the importance of having such a transparent database not only collaborating with educational institutions but also with other important sectors such as healthcare and social services that can benefit from this database [11]. In Ukraine, they particularly highlight the significance of elaborating the features of this database to ensure the system collects and processes statistical data on vulnerable children and is inclusive for children with disabilities because the current data processing does not provide such information [29].

One of the main problems in Group 3 countries is the unavailability of open data sources. In addition to this, different unethical data processing patterns emerge from the perspective of accuracy and reliability [16]. In Azerbaijan, there are discrepancies between two sets of data collected from official sources and data compiled independently by researchers or non-governmental organizations. Also, collecting all child data from the relevant government authorities is unavailable since unique identifiers are still not resolved. Additionally, data regarding out-of-school students are not differentiated in the public sources which means that identifying these students whether they dropped out or never attended school is very challenging and it overlaps [16]. In Georgia, we found similar tendencies where it was mentioned that data collection methods are complex enough and data related to the children out of education is limited [31]. According to the GEM report for Georgia, the data related to the students with limited abilities is traceable only if the student identifies themselves by registering with the Social Service Department [31]. In Armenia, there is a dedicated organization that collects data related to students with special needs, nevertheless, the challenge is schools' participation in data collection and use of data in policy development [21]. Schools in Armenia indicate a lack of interest in providing data input to the existing systems and it results in further difficulties in maintaining them.

Kazakhstan faces the same issues from the perspective of not possessing a unified and consistent database which raises doubts about the completeness of data on children with special needs [25]. Thus, the current data do not represent all learners in the country. Both the World Bank Report in 2018 and the GEM Report 2020 reported the same case in Uzbekistan the statistics on learners with special needs do not represent the real data [32]. In Kyrgyzstan, the data related to the learners with special needs are collected in the schools, and then they are sent to the Ministry of Education and Science [3]. The data is secured by providing individualized access for parents and students where they can log in through their credentials to update the system. The challenge happens in the data collection where they cannot keep tracking the students who drop out mainly due to major reasons that are immigration and teaching language. In Tajikistan, the data on students with special needs are collected from various authorities and for different purposes such as providing social support and planning the organization of the educational process [20].

Group 4 countries have their own systems where they process data. GEM Report for Russia mentions that in compliance with the Federal Law of the Russian Federation, most

Table 2: Ethical Data Challenges in Post-Soviet Countries*

Country	Accuracy	Bias	Fragmentation	Transparency
Armenia	•	•	•	
Azerbaijan	•	•		•
Belarus	•	•	•	•
Estonia	•			•
Georgia		•		•
Kazakhstan	•	•	•	
Kyrgyzstan	•	•		•
Latvia	•			
Lithuania	•		•	
Moldova		•	•	•
Russia	•	•		•
Tajikistan	•			•
Turkmenistan	•	•		•
Ukraine	•			•
Uzbekistan	•	•		•

* The symbol "•" indicates reported issues on the respective ethical data processing point.

educational organizations have the documentation on processing and protection of personal data [19]. This protects parents and students from the illegal use of their personal data. The main challenge in Russia is the openness of the data resources which cannot be accessed publicly and collected data is not valid due to overlapping issues. In Russia, schools provide data on students with special needs could also be a disabled student and vice versa. In the case of Belarus, the situation, most of this data can only be found on the website of the Ministry of Education but this only provides general statistics without going a deeper analysis [27]. Apart from this fragmented and basic available statistics on learners with special needs, there is no official data reporting on the pilot schools where inclusive education was implemented in Belarus. Moreover, the Ministry of Education in Belarus does not include individuals who have psychophysical development challenges to put under the category of children with special needs, thus these students are not represented in the database. Another problem is about the registration of students with special needs where they are represented in the national database only if their parents allow or self-identify themselves. Overall, the GEM Report for Belarus highlights that using general statistics may not always provide an accurate portrayal of how the educational process is structured for children with special needs. UN Turkmenistan reports similar data processing issues where most of the students with special needs are not registered and represented in the collected data [33]. Moreover, the detailed data across different sectors is not easily accessible.

4. CONCLUSION

This paper aims to present current challenges in post-Soviet countries from the perspective of ethical data processing. Even though the types of issues differ in each group of coun-

tries mentioned in Table 1, the most common challenges were about the representation of the students with special needs in the data collection process. It happens due to several reasons. Initially, most countries rely on parents to register their children in the database and if not registered, then these students are not considered to be learners with special needs. Secondly, the definition of special needs varies based on the country as well. While some countries categorize neurodivergent students as individuals with special needs, others do not acknowledge or classify them in the same manner.

This research can be extended in various ways. Firstly, future research can focus on analyzing the fairness of the data collection by implementing educational data mining techniques to see the credibility of overall data ethics. Because our current was based on the information provided by the international non-governmental organizations and government bodies. Secondly, future research can also focus on conducting interviews with government officials to understand their actions in ethical educational data processing and the use of this data in policy development.

5. ACKNOWLEDGMENTS

6. REFERENCES

[1] P. 2022. Pisa 2022 participants. https://www.oecd.org/pisa/aboutpisa/pisa-2022-participants.htm. Accessed: March 11, 2024.

[2] A. E. Abylkassymova. System modernization of general secondary education in the Republic of Kazakhstan. *Revista Tempos e Espaços Em Educação*, 13(32):1–17, 2020.

[3] L. M. Aleksandr Ivanov, Anastasia Kutepova. GEM report 2021: Kyrgyzstan. *Global Education Monitoring Report - UNESCO*, 2021.

[4] C. Alexopoulos, S. Saxena, M. Janssen, N. Rizun, M. Lnenicka, and R. Matheus. Why do open government data initiatives fail in developing countries? a root cause analysis of the most prevalent barriers and problems. *The Electronic Journal of Information Systems in Developing Countries*, 90:e12297, 2023.

[5] E. Amélie Lecheval. GEM report 2021: Lithuania. *Global Education Monitoring Report - UNESCO*, 2021.

[6] R. Anderson. Thematic content analysis (tca). *Descriptive presentation of qualitative data*, 3:1–4, 2007.

[7] A. Åslund, P. Boone, S. Johnson, S. Fischer, and B. W. Ickes. How to stabilize: Lessons from post-communist countries. *Brookings papers on economic activity*, 1996(1):217–313, 1996.

[8] G. Babayeva. Gender equality in the education system of azerbaijan republic. *Collection of scientific papers*, (July 8, 2022; Paris, France):261–263, 2022.

[9] L. Carlsen and R. Bruggemann. The 17 united nations' sustainable development goals: A status by 2020. *International Journal of Sustainable Development & World Ecology*, 29(3):219–229, 2022.

[10] S. R. Carroll, D. Rodriguez-Lonebear, and A. Martinez. Indigenous data governance: strategies from united states native nations. *Data Science Journal*, 18:18–31, 2019.

[11] V. R. for Association Educational Centre Pro Didactica. GEM report 2021: Moldova. *Global Education Monitoring Report - UNESCO*, 2021.

[12] D. Hernández-Torrano, L. Karabassova, Z. Izekenova, and M. G. Courtney. Mapping education research in post-soviet countries: A bibliometric analysis. *International Journal of Educational Development*, 87:102502, 2021.

[13] A. Karimov, M. Saarela, and T. Kärkkäinen. Clustering to define interview participants for analyzing student feedback: a case of legends of learning. In *International Conference on Educational Data Mining*, pages 234–243, 2023.

[14] A. Karimov, M. Saarela, and T. Kärkkäinen. The impact of online educational platform on students' motivation and grades: the case of khan academy in the under-resourced communities. In *International conference on educational data mining*, pages 466–473. International Educational Data Mining Society, 2023.

[15] A. Karimov, M. Saarela, and T. Kärkkäinen. Understanding teachers' perspectives on ethical concerns and skills to use ai tools. In *International Conference on Learning Analytics and Knowledge*, pages 230–232. Society for Learning Analytics Research (SoLAR), 2024.

[16] E. Kazimzde and U. Mikayilova. GEM report 2021: Azerbaijan. *Global Education Monitoring Report - UNESCO*, 2021.

[17] E. Marcella Turner-Cmuchal. GEM report 2021: Estonia. *Global Education Monitoring Report - UNESCO*, 2021.

[18] M. N. Meyer. Practical tips for ethical data sharing. *Advances in methods and practices in psychological science*, 1(1):131–144, 2018.

[19] M. Moiseeva. GEM report 2021: Russia. *Global Education Monitoring Report - UNESCO*, 2021.

[20] K. Mukhtori. GEM report 2021: Tajikistan. *Global Education Monitoring Report - UNESCO*, 2021.

[21] L. Nazaryan. GEM report 2021: Armenia. *Global Education Monitoring Report - UNESCO*, 2021.

[22] U. C. on the Rights of Persons with Disabilities (18th sess. : 2017 : Geneva). Concluding observations in relation to the initial report of latvia : Committee on the rights of persons with disabilities : draft / prepared by the committee. Agenda information CRPD/C/18/1 5 Consideration of reports submitted by parties to the Convention under article 35., Aug. 23 2017. 8 p.

[23] PISA. Pisa participants. https://www.oecd.org/pisa/aboutpisa/pisa-participants.htm. Accessed: March 11, 2024.

[24] H. N. Prasetyo and S. F. S. Gumilang. Data governance strategy for e-government in bandung district governments. *International Journal of Engineering & Technology*, 8(1.9):254–258, 2019.

[25] J. Ramberg. Global education monitoring report 2021–central andeastern europe, the caucasus and central asia–inclusion and education: All means all, 2021.

[26] E. Rustamov, U. Zalova-Nuriyeva, M. Allahverdiyeva, T. Abbasov, and N. Rustamova. Azerbaijani adaptation of the perceived school experience scale: examining its impact on psychological distress and school satisfaction. *Problems of Education in the 21st Century*, 81(6):869, 2023.

[27] H. Siarova. GEM report 2021: Belarus. *Global Education Monitoring Report - UNESCO*, 2021.

[28] I. Silova and S. Niyozov. *Globalization on the margins: Education and post-socialist transformations in Central Asia*. IAP, 2020.

[29] N. Sofiy. GEM report 2021: Ukraine. *Global Education Monitoring Report - UNESCO*, 2021.

[30] T. Suleymanov. Transformation of higher education in azerbaijan: reforms, policies and current trends. *J Econ Sci Theory Pract*, 77:40–60, 2020.

[31] M. Tamar and M. Giorgi. GEM report 2021: Georgia. *Global Education Monitoring Report - UNESCO*, 2021.

[32] J. Teixeira, J. Gresham, J. Liberman, K. Miyamoto, D. Chukmaitova, K. Gonzalez, I. Sadikova, I. Ciucanu, I. Ahadjonov, M. Pomes-Jimenez, M. Dinlemez, O. John, S. Vasudevan, S. Smid, and R. Miorelli. Uzbekistan education sector analysis final report, 12 2018.

[33] U. N. Turkmenistan. Situation analysis of the boys and girls with disabilities in turkmenistan. https://turkmenistan.un.org/en/835-situation-analysis-boys-and-girls-disabilities-turkmenistan. Accessed: March 11, 2024.

[34] UN-iLibrary. Global education monitoring report. https://www.un-ilibrary.org/content/periodicals/26180693. Accessed: March 11, 2024.

[35] E. Verity Donnelly. GEM report 2021: Latvia. *Global Education Monitoring Report - UNESCO*, 2021.

[36] J. Yebenes and M. Zorrilla. Towards a data governance framework for third generation platforms. *Procedia Computer Science*, 151:614–621, 2019.