Inclusive education evaluation by using Big Data sources of information

Descriptive Analysis

Final Report

Oxford Policy Management

Commissioned by UNICEF, Kazakhstan

August 2023

About Oxford Policy Management

Our vision is for fair public policy that benefits both people and the planet. Our purpose is to improve lives through sustainable policy change in low- and middle-income countries.

Through our global network of offices, we work in partnership with national stakeholders and decision makers to research, design, implement and evaluate impactful public policy. We work in all areas of economic and social policy and governance, including health, finance, education, climate change and public sector management. We have cross-cutting expertise in our dedicated teams of monitoring and evaluation, political economy analysis, statistics, and research methods specialists. We draw on our local and international sector experts to provide the very best evidence-based support.

Oxford Policy Management Limited
Registered in England: 3122495

Level 3, Clarendon House
52 Cornmarket Street
Oxford, OX1 3HJ
United Kingdom

Tel: +44 (0) 1865 207 300
Fax: +44 (0) 1865 207 301
Email: admin@opml.co.uk
Website: www.opml.co.uk
Twitter: @OPMglobal
Facebook: @OPMglobal
YouTube: @OPMglobal
LinkedIn: @OPMglobal

Preface

This study was carried out by Oxford Policy Management (OPM) between January and July 2023. The project manager was Udayan Rathore and the remaining team members included Ida Brzezinska (Data Analyst), Arseniy Gurin (Senior Data Scientist), Paul Jasper (Team Leader), Elayn Sammon (Senior Inclusion Policy Expert), and Zhanar Zhaxylykova (National Evaluation Expert). For further information, contact Udayan Rathore at Udayan.rathore@opml.co.uk.

The contact point for the client is Udayan Rathore, who can be reached out to at Udayan.rathore@opml.co.uk. The client reference number for the project is A5421.

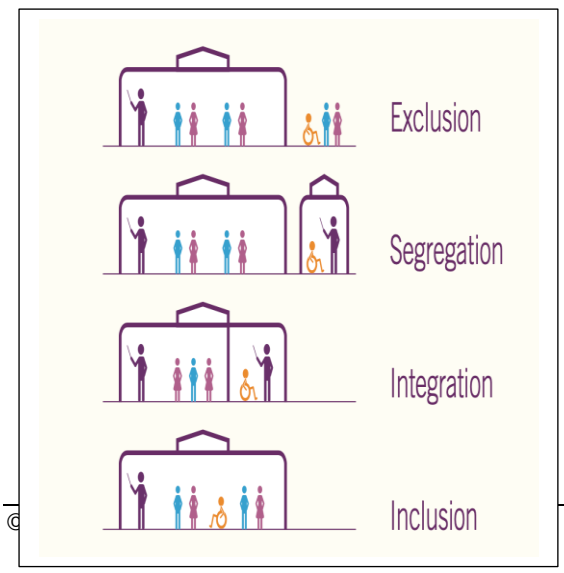
The implementation of this study, 'Inclusive education evaluation by using Big Data sources', would not have been possible without extensive contributions from a range of organisations and people. We are grateful to the officials/experts of UNICEF Kazakhstan for trusting us with this study and providing the financial support to undertake our research. We would like to thank Tatiana Aderkhina and Raushan Ibrasheva for their input at the inception phase of the project, as well as for their comments on the preliminary findings presentation. We are also thankful to the Junction Bulgaria team, whose work on the formative evaluation provided us with valuable insights on the inclusive education landscape in Kazakhstan. We would also like to express our gratitude to the reviewers of the inception phase presentation at UNICEF.

We are grateful to all the individuals and organisations who posted their perspectives on inclusive education in Kazakhstan on publicly accessible social media platforms, without which this research would not have been possible.

We would also like to acknowledge the contributions of OPM staff members Varun Vashistha and Vladimir Imnaishvili in managing the administrative and financial aspects of this study. We would also like to thank Matilda Eriksson for assisting us with operational details of Meltwater, which was used for analysing one component of social media data in the study.

All opinions expressed, and any mistakes in the report, remain the responsibility of the OPM team.

Figure 1: Conceptual difference between exclusion, segregation, integration, and inclusion



Exclusion: No recognition of the right or capacity of children with disabilities to education, and consequent denial of access to education in any form.

Segregation: Placement of children with disabilities in separate environments designed or used to respond to particular or various impairments, in isolation from children without disabilities.

Integration: Placement of children with disabilities in existing mainstream educational institutions, as long as the child can adjust to fit the standardised requirements of such institutions.

Inclusion: A process of systemic reform embodying changes and modifications in content, teaching methods, approaches, structures, and strategies in education to overcome barriers. It requires a commitment to changing the system to fit the student.

Source: UNESCO, UNICEF, and FCDO (2021), p. 35

We also pay attention to measures that are not specific to inclusive education, but which can contribute to an enabling environment for inclusive education, such as measures to prevent disability, for early identification, and to protect children from violence and abuse in the education system.

This evaluation is important for a variety of reasons. First, it complements a formative evaluation commissioned by UNICEF that is ongoing at the time of writing and that seeks to assess how state education programmes have contributed to improving the inclusive education landscape in Kazakhstan. We refer to this parallel formative evaluation as the 'Junction Bulgaria study'. Second, instead of focusing on specific programmes, this study uses innovative analytical methods to compile and synthesise publicly accessible social media and online news information from two key stakeholder groups, whose perspectives are categorised into two text corpora and analysed. The first group of stakeholders comprises of the general online public in Kazakhstan, and includes parents, teachers, and youth overall. The second stakeholder group caters to the views and institutional knowledge of organisations active in inclusive education activities in Kazakhstan. Third, the use of non-traditional data and analytical methods provides a space for experimentation and engagement with different viewpoints, some of which may be difficult to observe and are thus often left out of traditional methods and approaches.

Executive summary

Introduction

To better understand the inclusive education landscape in Kazakhstan, the United Nations Children's Fund (UNICEF) commissioned the present study with the objective of understanding different stakeholder perspectives as expressed via social media and online news media, using social media listening (SML) tools. This research is expected to generate insights for UNICEF and the Ministry of Education, Government of Kazakhstan, to inform future strategies and policies in this area. The findings provide insights for leading scientific institutions in Kazakhstan actively working in the field of inclusive education. Other development partners such as the United Nations Educational, Scientific and Cultural Organization (UNESCO), the United Nations Development Programme (UNDP), and the World Bank, as well as multiple other non-governmental organisations (NGOs) working and supporting investments in education in Kazakhstan, are also likely to benefit from these findings.

Evaluation purpose, objectives, and scope

The primary research objective is to use SML tools and quantitative text analysis tools to understand the perspectives of individuals and organisations, as expressed in text online, on inclusive education in Kazakhstan. By quantitatively analysing non-traditional online news and social media sources of data, we present insights and formulate recommendations based on viewpoints from different stakeholders that might otherwise be difficult to observe. Our research complements an ongoing formative evaluation (the Junction Bulgaria study).

This evaluation aims to answer three evaluation questions (EQs) for both stakeholder groups.

1. **What is the content of social and online media discussions on inclusive education in Kazakhstan?** We interpret 'content' to relate to what these discussions 'look like', i.e. what is being said, where, by whom, and when. Note that we purposefully only deploy quantitative text analysis techniques to answer this question. This means that descriptions of this 'content' rely on quantitative analytical techniques.

2. **What do people mention as factors 'holding back' (barriers) inclusive education?** We interpret 'barriers' broadly, in the sense that we try to find out what is being mentioned online when problems and issues with improving access and promoting the **rights** of children to inclusive education are discussed. As above, we deploy quantitative text analysis techniques to answer this EQ.

3. **Do people mention state programmes at all?** Here, we search for a predefined list of state programmes.

This study was designed to be formative and descriptive in nature and is thus a part of a broader formative evaluation exercise. The focus is on describing online news and discussions on inclusive education in Kazakhstan rather than on evaluating an 'object'.

In the inception phase, detailed discussions were undertaken with UNICEF to define and operationalise 'inclusive education' in the context of this study. Based on these discussions, *children under 18 years of age with disabilities* were identified as the target

group of interest. This includes children with long-term physical, mental, intellectual, or sensory impairments, which in interaction with various barriers may hinder their effective participation in society on an equal basis with others. This approach integrates the medical and social model of disability in our conceptualisation of disability (UNESCO, UNICEF, & FCDO, 2021).

A vital component of operationalising inclusion relates to the level of education where this inclusion occurs or is desired. In this study, we focus on children in preschool as well as those receiving primary and secondary education. Due to multiple considerations, as mentioned later, we also limit our study to the relatively recent past, from the beginning of 2022 until May 2023, where our data is restricted to cover conversations and posts online captured by our online text scrapers.

Methodology

To study the EQs from different perspectives and triangulate insights, we focused on two primary stakeholder groups, whose publicly accessible social media data were organised into two different text corpora (collections of text data). We summarise each below.

1. **Text data in Corpus 1** represents discussions of the general online public. This was created using Meltwater, an online platform which offers scraping and analytics of social media and online news data.

As a first step in building a search query for Corpus 1 in Meltwater, we identified multiple subtopics within the broad topic of inclusive education through a literature review and expert discussions. For each subtopic, we created a combination of keywords in Russian and Kazakh that accurately described the chosen topic and considered the specifics of each language and the different groups of interest. Once our subtopics and keywords were defined, we initiated search queries, scraping online discussions on inclusive education in Kazakhstan. In our search queries, we limited the geography to Kazakhstan by including a 'Location' query that specified our results must mention locations within Kazakhstan. To ensure the quality of this search, we repeatedly validated queries by human review.

2. Text data from accounts and websites of specific organisations active in the inclusive education space in Kazakhstan were collected as part of **Corpus 2** and built using self-programmed scrapers.

Through a purposive research design, we identified multiple organisations and their publicly accessible social media handles and websites, which were scraped via a self-written programme using Python that downloaded publicly accessible text. A major component of Corpus 2 came from Instagram channels, Telegram groups and channels, and web sources such as blogs and news outlets.

Findings

For **EQ 1**, we summarise key characteristics of the social and online media discussions on inclusive education in Kazakhstan as follows.

1. **Inclusive education in Kazakhstan is a topic of ongoing discussion, albeit a small one compared to general discussions on education online.** Analysis of Corpus 1 and Corpus 2 shows an overall increasing trend in discussions on inclusive education.

However, we estimate that inclusive education discussions in Corpus 1 make up only about 3–6% of all online text on education more broadly.

2. **The general public's online engagement (Corpus 1) with inclusive education is higher on news sites, blogs, and Facebook** than on other social media sites. Organisations working on inclusive education (Corpus 2), however, mostly use platforms like **Instagram** and **Telegram**.

3. **Analysis of top news by reach in Corpus 1 shows that the Government of Kazakhstan is successfully promoting its active measures to develop inclusive education** with the introduction of modern technologies, the creation of feedback channels, and initiatives aimed at protecting the rights of children with special educational needs (SEN). This shows that the government's messaging has a wide reach, and that they are successfully highlighting and promoting these activities.

4. **Spatial patterns of online conversations from Corpus 2 show that these discussions are mostly concentrated in large urban cities** like Astana and Almaty, even after accounting for their higher population densities. This is suggestive of relatively higher engagement on inclusive education in these areas.

5. **Negative news in Corpus 1 often focused on the lack of protection of the rights of children with SEN, as well as on the quality of their education.** Meanwhile, positive news highlighted measures taken to support **children with autism**, to prevent **disability in newborns**, to increase the **capacity building of teachers**, and for **social protection**.

6. **A hashtag analysis in Corpus 1 shows that autism and cerebral palsy are the most widely discussed disabilities.** Also, the **adoption of testing for children with SEN is actively discussed,** as indicated by the United National Testing (EHT) hashtag. Among non-profit organisations, work of the Bulat Utemuratov Foundation has received attention and positive discussion for its contribution to education for children with autism.

7. Analysis of the most frequently used words from Corpus 2 shows discussions emphasised the **capacity building of teachers, specialists and measures to support inclusive education.** As before, **autism** was the most discussed disability.

8. **Topic analysis on Corpus 2 identified the most common themes of discussions, and again features the capacity building of teachers, school education content and pedagogy** as prominent topics. Simultaneously, early diagnosis and treatment, social support and opportunities for children, were also vital themes of discussion.

For **EQ 2**, our key findings are as follows.

9. **An analysis of 'Rights and Barriers' from Corpus 1 reveals that incidents of abuse, violence, and bullying in kindergartens, schools, universities, and families are being mentioned as *major* issues affecting adequate access to inclusive education.** We understand that these have been identified by the Government of Kazakhstan, which has made efforts to address quality issues, improve standards, and establish minimum safety requirements in early childhood education.

10. **In Corpus 2, the need for improved teacher training and the capacity building of teachers and specialists who can support the SEN of children** feature prominently as major barriers to be overcome.

11. **In discussions on inclusive education online, girls are significantly underrepresented.** With an eye on assessing differences in how inclusive education discussions make references to boys and girls, we check if girls or boys are disproportionately represented in discussions on inclusive education when compared to discussions about education with no mention of disability. **We find that the share of messages that mention boys and girls across the corpora of 'education and disability' and 'education' is comparable.** However, the share of messages that mention boys in such discussions is about three times larger than girls and accounts for about 75% of the overall discussions. This suggests that girls, and possibly their specific needs, are underrepresented in such discussions.

For **EQ 3**, our key finding is as follows.

12. **We find limited mentions of specific programmes on inclusive education in Corpus 1 and Corpus 2.** In the latter, we do find mentions of 'Educated Nation' (Образованная нация) and 'Comfort School' (Комфортная школа). This is not to say that these programmes are not important, but it indicates that specific issues, rather than government programmes, are being discussed more widely on social media. It may also imply that the government is not able to promote programmes in a way that they are picked up in social media discourse around inclusive education.

Key policy recommendations

1. **Timing and choice of broadcasting platform for key messages:** We find that engagement with inclusive education rises with the start of the academic year and tapers off gradually. It is sensitive to initiatives such as the month of raising awareness about autism. Similarly, general public's online engagement is higher on news portals and blogs, whereas social media engagement on Instagram and Telegram is higher for organisations working on inclusive education. This information can be leveraged by the Government of Kazakhstan, particularly by the Ministry of Education and the Ministry of Information and Communications, to relay key messages on inclusive education using varied platforms for different stakeholders. In this context, evidence shows positive effects of such media messaging on children's education (Keefer & Khemani, 2014).

2. **Focus on enhancing the accessibility of preschool, primary, and secondary education facilities via accessibility audits:** Despite positive perceptions around the increased investment on upgrading school infrastructure and introducing modern technologies, the Junction Bulgaria team found that there is a requirement for even higher investment in physical infrastructure. Accessibility audits can help prioritise necessary modifications and ensure compliance with accessibility standards to enhance physical access to educational facilities. Certain relevant voluntary initiatives have already emerged—notably Accessible Kazakhstan—and could provide guidance for the Government of Kazakhstan to institutionalise this model for educational facilities.

3. **Investment in human capital:** Our findings across corpora indicate a perception of lack of capacity building for teachers. Provision of training and professional development opportunities for teachers and staff members on inclusive education practices, disability awareness, and accessibility guidelines will enable educators to better support students with disabilities and create a more inclusive learning environment. It could also contribute towards a more individual approach being taken to students with SEN.

4. **Coverage of social media dark areas:** The social data media data on which this analysis is based comes from larger cities like Astana and Almaty. However, more research is needed to understand perceptions of, and barriers to, inclusive education in other parts of the country.

5. **Gender-sensitive assessment of educational needs:** Our results show that girls are relatively underrepresented in online discussions on inclusive education in Kazakhstan, suggesting a lack of attention paid to the differential needs of girls in the educational system. This calls for a collection of indicators disaggregated by gender within the education sector in order to identify gender-specific factors limiting access to inclusive education. It also calls for a general focus on gender-sensitive messaging on the needs of girls in inclusive education.

6. **Early childhood interventions:** Early diagnosis and treatment of disability is a recurring topic in our corpora. Early prevention and identification of disabilities, as well as provision of services for early childhood development for children with SEN, are crucial to ensure their educational needs are addressed from the outset. Positive news online indicates some progress has been made in measures taken to prevent disability in newborns and young children, including through better equipping of perinatal centres. These initiatives should be expanded in coverage and coupled with correctional support to prepare young children with SEN to study in public schools.

7. **Adaptation of standardised testing:** The EHT hashtag appeared commonly in our analysis, showing active discussions around the need to adapt standardised tests to the needs of children with SEN. Our findings from top news in Corpus 1 indicate that the Government of Kazakhstan has provided an additional 40 minutes for children with SEN during such tests, showing that progress on this issue is already underway. Further adjustments could be made to adapt testing materials, for instance for children who cannot see the text.

Key recommendations for future use of SML in evaluations

1. **Designing EQs based on what SML can and cannot answer:** As against traditional impact evaluations that are focussed on causal impact of an intervention, SML is better suited to address descriptive questions that are intended to capture the perspectives of stakeholders on a specific topic of online discussion.

2. **Validation of online data:** Building search queries should be iterative and based on expert discussions. The list of keywords should account for linguistic or cultural sensitivities, while being simple enough to ensure a sizeable text corpus for analysis.

3. **Accounting for exclusion and inclusion bias:** Inequalities in access to social media may exclude certain groups from online discussions, thereby limiting the representativeness at the national, or regional level. Bias from extreme views on social media than in person also highlights the need for expert discussions to validate the relevance of online posts.

4. **External Validity:** As online discussions are not necessarily representative of the entire population and in some cases the number of posts is small, the evaluation should clearly state whether the results can be generalisable.

Table of contents

Preface i

Executive summary iii

List of tables, figures, and boxes ix

List of abbreviations xi

1 Introduction 1

1.1 The object of the evaluation and defining 'inclusive education' 1

1.2 Inclusive education in Kazakhstan 3

1.3 Evaluation stakeholders 4

1.4 Evaluation management and governance 6

2 Evaluation objectives, framework, and EQs 9

2.1 Purpose and objectives 9

2.2 Policymaking cycle in Kazakhstan 9

2.3 Groups of interest in this present study 10

2.4 Evaluation scope 11

2.5 EQs 12

3 Methods 14

3.1 Scraping and processing text data from the internet 14

3.2 Data analysis 18

3.3 Ethics 20

3.4 Human rights, gender, and equity 20

3.5 Stakeholder participation 21

3.6 Methodological limitations 21

4 Findings 23

4.1 Corpus 1: discussions by the general public on inclusive education online 23

4.2 Corpus 2: Online discussions by organisations focusing on inclusive education in Kazakhstan 40

4.3 Our findings in the context of the broader formative evaluation and implications for inclusive education policymaking in Kazakhstan 48

4.4 Caveats, gaps, and limitations 49

5 Conclusions and recommendations 51

5.1 Conclusions 51

5.2 Recommendations regarding inclusive education in Kazakhstan 53

5.3 Recommendations for the future use of SML in evaluations 55

6 References 57

Annex A Terms of reference 58

Annex B: Additional Figures and Tables 61

List of tables, figures, and boxes

Table 1: Analytical tools used for each EQ and stakeholder group 13
Table 2: Number of results from key search queries 16
Table 3: Number of results from key search queries 18
Table 4: Topic modelling (English) for 7 topic identified by the algorithm 43
Table 5: Topic modelling of Rights Barriers in corpus 2 (English) 45
Table 6: Corpus 2, List of Organisations 61
Table 7: Topic modelling (Russian) 63
Table 8: Topic modelling of Rights Barriers in corpus 2 (Russian) 64

Figure 1: Conceptual difference between exclusion, segregation, integration and inclusion ... i
Figure 2: Sub-topics within Inclusive Education 14
Figure 3: Key sub-topics in our final search queries 16
Figure 4: Mentions trends—Queries on Inclusive Education 23
Figure 5: Top publications by reach—Queries on Inclusive Education 24
Figure 6: Top news—Query 1: Inclusive Education KZ 24
Figure 7: Top news—Query 2: Inclusive Education RU 26
Figure 8: Top sources - Queries on Inclusive Education 26
Figure 9: Top negative news with the largest audience reach—Query 2: Inclusive Education RU 27
Figure 10: Top positive news with the largest audience reach—Query 2: Inclusive Education RU 28
Figure 11: Top positive news with the largest audience reach—Query 1: Inclusive Education KZ 28
Figure 12: Sentiment analysis 29
Figure 13: Hashtag analysis—Instagram and Facebook 30
Figure 14: Top keywords 31
Figure 15: Mentions trend - Rights Barriers 32
Figure 16: Top News - Rights Barriers 33
Figure 17: Top negative news with the largest audience reach—Query 3: Rights Barriers RU 34
Figure 18: Top positive news with the largest audience reach—Query 3: Rights Barriers RU 34
Figure 19: Top news with the largest audience reach—Query 4: Rights Barriers KZ 35
Figure 20: Top keywords (Rights Barriers) 35
Figure 21: Gender analysis (Corpus 1) 36

Figure 22: Differences in sentiments across ‘education’ and ‘education and disability’, by gender (Corpus 1) 38
Figure 23: Spatial concentration of messages on inclusive education 40
Figure 24: Mentions trends for Corpus 2 41
Figure 25: Most common discussion words 42
Figure 26: Messages per topic 43
Figure 27: Topic prevalence over time, by month 44
Figure 28: Mention of Rights Barriers (Corpus 2) 44
Figure 29: Topics of Rights Barriers (Corpus 2) 45
Figure 30: Most cited words on barriers (Corpus 2) 46
Figure 31: Gender analysis (Corpus 2) 47
Figure 32: Building a search query in Meltwater (KZ) 61
Figure 33: Building a search query in Meltwater (RU) 61
Figure 34: Theory of Change (Junction Bulgaria) 66

Box 1: An example of the medical versus social model of disability 2
Box 2: Key words for Sub-topic 2: Physical disabilities 15
Box 3: Examples of news items on autism and cerebral palsy 30
Box 4: Additional key words associated with ‘Rights and Barriers domain’ 32

List of abbreviations

ASD	Autism Spectrum Disorder
EHT	United National Testing
EQ	Evaluation Question
ICT	Information and Communication Technology
NGO	Non-Governmental Organisation
NLP	Natural Language Processing
OPM	Oxford Policy Management
QA	Quality Assurance
SEN	Special Educational Needs
SML	Social Media Listening
ToC	Theory of Change
ToR	Terms of Reference
UNDP	United Nations Development Programme
UNESCO	United Nations Educational, Scientific and Cultural Organization
UNICEF	United Nations Children's Fund

1 Introduction

Kazakhstan is a relatively young country: close to 34% of its population is below the age of 18, of which over 188,000 children have SEN, a number that is growing every year (Ministry of Education, 2023).1 In 2020, the total number of children with special educational needs identified by the Psychological, Medical and Pedagogical Consultation (PMPK) stood at about 161,826 individuals, which accounts for about 2.6% of overall children population. The share of children with SEN is particularly high in age groups of 7–10 years (32.7%), 3–6 years (30%), and 11–15 years (26.6%). In terms of the geographical distribution, the highest number of children with SEN resided in the Turkestan region (12%), followed by the East Kazakhstan region (9.6%), Almaty region (9%), and Karaganda region (8%), respectively (Ministry of Education of the Republic of Kazakhstan, 2021).[2] To better understand the inclusive education landscape in the country, UNICEF commissioned the present innovative study with the objective of understanding different stakeholder perspectives on inclusive education using SML tools.[3]

The analysis aims to assess and identify perspectives, discussions, and opinions on access to inclusive education in the country. The innovative approach gives a voice to stakeholders who may otherwise be left out of traditional survey-based studies. Importantly, with close to 61% of its population having presence on different social media platforms,[4] Big Data research using organically generated social media data provides a unique opportunity to conduct rigorous research and support conventional evaluation approaches that use in-person data collection methods. The study generates insights that could be used by UNICEF and the Government of Kazakhstan to identify existing gaps in the inclusive education landscape and provide evidence to inform future strategies and work in this area. The present study complements a separate formative evaluation undertaken by the Junction Bulgaria team to assess how state education programmes have contributed to improving the inclusive education landscape.

1.1 The object of the evaluation and defining ‘inclusive education’

Unlike other evaluation exercises, which evaluate an ‘object’ such as a programme intervention, this study was designed to be a formative and descriptive evaluation product and is therefore a part of a broader formative evaluation exercise. The focus here is on describing the broader topic of online news and discussions of inclusive education in Kazakhstan, rather than to conduct an evaluation of an ‘object’. Importantly, and as agreed during the contracting phase, this study does not have a conventional Evaluation Matrix or a

[1] Concept Note No. 249, Ministry of Education, Government of the Republic of Kazakhstan, 28 March 2023: www.gov.kz/memleket/entities/edu/documents/details/451747?lang=ru (accessed 15 July 2023).

[2] According to Kazakhstan's legislature, children with disabilities receive financial assistance from the government. According to data from the National Statistics Bureau, the number of children with disabilities receiving such assistance as of 2021 was 98,254 https://bala.stat.gov.kz/chislennost-detej-invalidov-poluchayuschih-sotsialnye-posobiya-i-srednij-razmer-gosudarstvennyh-sotsialnyh-posobij/ (accessed 30 August 2023) .

[3] At an earlier stage, this project also had a spatial data analysis component to disaggregate regional variation in inclusive education outcomes across different levels of schooling. During the contracting phase, UNICEF and OPM together decided to restrict the scope of the study and exclude this component.

[4] https://datareportal.com/reports/digital-2023-kazakhstan (accessed 11 July 2023).

Theory of Change (ToC). This is because we do not explicitly test the validity of a ToC within this study. However, a draft ToC applicable to the formative evaluation was made available at the inception phase and can be found in the corresponding Annex (Figure 34). Importantly, the study design ensures that the EQs (Section 3) are mapped to the objectives of the study, as mentioned in the Terms of Reference (ToR) (Annex A), via rigorous methodological tools that are consistent with the relevant literature on SML.

During the inception phase of the project, detailed discussions were undertaken with UNICEF to define and operationalise 'inclusive education' in the context of this study. Based on these discussions, children with disabilities below 18 years of age were identified as the target group of interest. Children with disabilities include those who have long-term physical, mental, intellectual, or sensory impairments, which in interaction with various barriers may hinder their full and effective participation in society on an equal basis with others.[5] This approach integrates the medical model and social model of disability in our conceptualisation of disability (UNESCO, UNICEF, & FCDO, 2021). While the medical model is based on the characteristics of diseases and disorders, the social model acknowledges that disability is multidimensional and that the complex interaction of health and contextual factors may enable or restrict children from participating in education and learning effectively (Box 1).

Box 1: An example of the medical model versus the social model of disability

Aigul is a secondary school student whose eyesight is deteriorating because of diabetic retinopathy, a common complication of diabetes. The medical model sends her to a clinic, but there is no treatment to halt the progressive damage to her sight. It becomes more difficult for her to get to school using public transport, and her parents are trying to find extra money to afford taxis. Aigul is excluded because she cannot read texts or afford the extra transport costs. She is very worried that she will have to drop out of school.

The social model finds out what assistance she needs to continue attending school and to keep up with her studies. This means finding solutions to overcome the barriers to her inclusion. For example, her teachers and the Department of Education could lobby IT developers to make scanner apps for blind and visually impaired users available in her language. At the same time, the occupational therapist can work with Aigul, her teachers, and classmates to find, and practice with her, the most practical route to school by public transport. Aigul can also apply for a disability allowance to cover the additional disability-related costs.

Aigul has an impairment (deteriorating eyesight) which limits her functioning (reading textbooks and using public transport) until the participation restriction is lifted (by having access to the state-of-the-art scanner apps and access to support to learn new skills for independent travel).

An important component of operationalising inclusion relates to the level at which this inclusion occurs or is desired. We consider children in preschool, as well as those receiving primary and secondary education. Importantly, inclusion here means more than 'just' being physically present at school and recognises there are important distinctions between exclusion, segregation, integration, and inclusion. These terms are discussed below and outlined in Figure 1. This overall conceptualisation of inclusion informs our interpretation of results in Section 4.

[5] United Nations Convention on the Rights of Persons with Disabilities Article 1: https://social.desa.un.org/issues/disability/crpd/article-1-purpose.

1.2 Inclusive education in Kazakhstan

Inclusive education in Kazakhstan is an evolving process guided by several national programmes and projects.

According to the 2007 Law on Education, inclusive education is defined as the process to ensure equitable access to appropriate educational training for students with special individual abilities. The 2015 Conceptual Approaches to the Development of Inclusive Education in the Republic of Kazakhstan, a Policy Note endorsed by the Ministry of Education, further elaborates that inclusive education is a transformative undertaking within the education system, grounded in the principle of equal rights to quality education for all individuals.

In 2022, the categories of children with SEN were extended via Decree 6 of the Ministry of Education to three categories: first, children with behavioural and emotional problems and unfavourable psychological factors; second, children with barriers of a socio-psychological, economic, linguistic, and cultural nature; and third, children with developmental disabilities. In this study, we focus on the third component. The first two categories are included only when they are related to mentions of disability and education.

Further, the development of inclusive education in Kazakhstan is guided by a series of national programmes and projects. The State Programme for the Development of Education in the Republic of Kazakhstan (2011–20) and the State Programme for the Development of Education and Science (2016–19) have played significant roles in shaping educational policies and initiatives. These programmes were followed by the State Programme for the Development of Education and Science (2020–25), which transitioned to the national project 'Educated Nation' in 2021. The Concept of Education Development of the Republic of Kazakhstan (2022–26) provides a strategic framework for future educational reforms at all levels of education.

Despite these efforts, challenges persist in fully implementing inclusive education in Kazakhstan. A shortage of qualified personnel and inadequate infrastructure in the education sector pose significant barriers (Junction Bulgaria, 2023). Additionally, attitudes and misconceptions about disabilities need to be addressed to foster an inclusive culture within the society (UNICEF, 2019; Allan & Omarova, 2021). Progress in the education sector has also stalled as a result of the COVID-19 pandemic, and children with disabilities are some of the most affected by sectoral challenges (Marteau, 2020). Progress towards achieving Sustainable Development Goal 4, Quality Education, is off-track regarding the participation rate in pre-primary organised learning and in terms of the net primary enrolment rate (Sustainable Development Report, 2023). To tackle these challenges, the government, educational institutions, and NGOs are working on allocating more resources, improving

infrastructure, and promoting awareness about the importance of inclusive education (Junction Bulgaria, 2023).

The findings from our SML research (Section 4) align with the context of inclusive education in Kazakhstan. As we will see in subsequent sections, our research indicates an overall increasing trend in discussions about inclusive education throughout the academic year (schooling calendar). The government's active measures to develop inclusive education, including the use of modern technologies, feedback channels, and initiatives to protect the rights of children with SEN, have been recognised and discussed on various online platforms used by the general online public.

Negative news coverage often focus on the lack of protection of the rights of children with SEN and concerns about the quality of their education. Conversely, positive news stories highlight measures taken to support children with autism, prevent disability in newborns, and improve inclusive education in Kazakhstan, including the capacity building of teachers and the introduction of social protection initiatives. Among types of disability, autism and cerebral palsy are the most widely discussed in the context of inclusive education.

According to our hashtag analysis, NGOs receive attention and positive discussion as a result of their contributions to education for children with autism. Discussions also emphasise the importance of addressing barriers to inclusive education, including the need for dedicated centres, protection of rights, and improved teacher training and pedagogy.

The research findings from analysing the perceptions of organisations working in the inclusive education space highlight that these discussions are mostly concentrated in large urban cities like Astana and Almaty. While social media engagement varies across regions, this concentration suggests higher engagement in these urban areas. The findings also indicate a gradual increase in discussions involving organisations working in inclusive education since 2022, with specific platforms like Instagram and Telegram groups being utilised for public discussions by these organisations.

In conclusion, while challenges persist, the findings from our SML exercise research demonstrate a growing awareness of, and engagement with, inclusive education. These findings reinforce the importance of addressing barriers, promoting capacity building for teachers, protecting the rights of children with SEN, and fostering a more inclusive culture. The commitment of the government, in collaboration with various stakeholders, is essential in advancing inclusive practices, supporting teachers, and providing the necessary resources to create inclusive learning environments in Kazakhstan.

1.3 Evaluation stakeholders

This evaluation provides valuable evidence for several key stakeholders involved with issues of inclusive education, as follows.

Ministry of Education: The Ministry of Education is the main government body responsible for developing and implementing education policies, including inclusive education. They develop regulatory acts that define strategic priorities and standards for inclusive education and coordinate the activities of other educational units.

National Academy of Education named after I. Altynsarin: This leading educational and scientific institution in Kazakhstan is actively involved in the development and implementation of educational programmes and research in the field of inclusive education.

The Academy organises scientific conferences, seminars, and training events aimed at enhancing the competency of teachers and specialists in inclusive education.

National Scientific and Practical Centre for Special and Inclusive Education Development: The centre is a leading scientific and methodological organisation involved in the development and dissemination of practical recommendations and methodologies in the field of inclusive education. It provides educational resources, conducts research, and offers advisory and methodological support to teachers and specialists in inclusive education.

The Office of Psychological and Pedagogical Correction: This is an institution where any child with SEN and his/her family can receive qualified assistance from specialists such as a speech pathologists and therapists, psychologists, typhlopedagogues, social pedagogues, physical therapy instructors, and music rhythm teachers. There are the offices all over Kazakhstan, funded by the government.

Specialists involved in psychological, medical, and pedagogical consultation: One of the types of special educational institution examines students for developmental disabilities and develops recommendations on education and training for parents and teachers. The main task of the consultation is to identify a child's capabilities and developmental disorders, as well as to determine the educational programme and special conditions for education in accordance with the individual needs of the child.

UNICEF: UNICEF in Kazakhstan actively supports the development of inclusive education. The organisation collaborates with the government and other partners, providing financial and technical support to projects related to inclusive education. UNICEF monitors and evaluates the implementation of programmes in this area.

Other development partners: UNESCO, UNDP, and the World Bank are among several development partners supporting investments in education in Kazakhstan that will also benefit from the research findings.

NGOs: NGOs, especially those representing parents and caregivers of children with disabilities, play a crucial role in supporting and advocating for the rights of children with SEN and disabilities. They monitor the implementation of inclusive education, provide consultation and support to parents, and participate in projects aimed at improving the quality of, and accessibility to, inclusive education, as well as the social integration of children into the society.

Among the organisations that operate in the inclusive education area in Kazakhstan, the following are key non-profits.

DARA Foundation: The Foundation actively works towards creating equal opportunities for all learners, including children with SEN. The Foundation conducts training programmes and workshops for educators, develops specialised teaching materials, and ensures access to adapted learning materials and technologies. It also actively supports the implementation of inclusive practices in schools and contributes to the creation of a friendly and barrier-free educational environment. The Foundation has initiated a system of support for inclusive education in general education schools through a network of Inclusive Education Resource Centres in almost all regions of Kazakhstan. These centres are based inside schools and have been created to help teachers implement the principles of inclusion in the learning process, as well as to create a barrier-free and friendly classroom and school environment for children with SEN.

Bolashak Corporate Foundation: The Foundation plays an instrumental role in promoting equal access to education for all individuals, including those with diverse needs and backgrounds. The Foundation actively supports initiatives that foster inclusive practices within educational institutions, such as providing resources, training programmes, and infrastructure enhancements. Through its partnerships and collaborations with the government, business, NGOs, and parents, the Bolashak Corporate Foundation helps create an inclusive educational environment that empowers students to thrive and reach their full potential.

Parents' community: Parents are important stakeholders in the field of inclusive education. They play an active role in advocating for their children's right to equal access to education and participate in the related decision making processes. Parents provide support to their children, collaborate with teachers and school administration, and engage with government bodies and NGOs.

Business: One of the largest quasi-state donor organisations is the **Samruk-Kazyna National Welfare Fund** of Kazakhstan that provides annual funding to charitable and social initiatives, including the improvement of inclusive education via its Samruk-Kazyna Trust Corporate Fund and the Qazaqstan Halqyna Public Foundation.

There are also several NGOs in Kazakhstan run by businesses. Among them is the **Bulat Utemuratov Foundation**, which endorses 11 'Asyl Miras' autism centres across Kazakhstan. Their mission is to recognise challenges and address gaps in social, communicative, and practical abilities among children with autism spectrum disorder (ASD) under the age of 15. The Foundation has established a resource centre, which engages in ongoing training and development in the realm of ASD child interaction. This extends to specialists at the Asyl Miras centres, NGO employees, medical institutions, and state correctional institutions, with roles as described above.

We describe the role of specific stakeholders in our evaluation in the next section, and describe stakeholder participation in Section 3.5.

1.4 Evaluation management and governance

On behalf of the Government of Kazakhstan, UNICEF Kazakhstan contracted the project, entitled 'Inclusive education evaluation by using Big Data source of information (descriptive analytics)', to Oxford Policy Management (OPM) on 29 December 2022.

The key deliverables, as per the ToR (Annex A), included:

1. the inception report, delivered to UNICEF as a PowerPoint presentation on 12 March 2023 (a revised version incorporating suggestions was shared on 28 March 2023);
2. the present final descriptive report (called the digital ecosystem analysis report in the ToR), which serves as the main technical report and includes results from all key analyses; and
3. the final 'package', including the final methodological note and tools.

OPM assembled an exceptional team of experts to implement this assignment. The team structure was as follows.

- Paul Jasper was the team leader. Given his extensive experience with innovative data approaches in evaluations, including geospatial analysis and text analytics, he provided technical leadership to the rest of the team.
- The data science team was composed of three key members:
    - Arseniy Gurin, the Senior Data Scientist, who provided technical data science and data engineering oversight; and
    - Udayan Rathore and Ida Brzezinska, quantitative data analysts, who implemented most of the data processing and analysis (Udayan was also the project manager).
- Zhanar Zhaxylykova was the National Evaluation Expert. As a Kazakh national and experienced education consultant, her input ensured that the evaluation was firmly grounded in the local inclusive education context. Zhanar is based in Kazakhstan and also served as our local contact point.
- Elayn Sammon, was our Senior Inclusion Policy Expert. Her input ensured that evaluation objectives and findings were relevant to the promotion of inclusion in education in Kazakhstan.

The overall oversight of the contract sat with the UNICEF Kazakhstan Team, as the main client and holder of the contract with OPM. Regular meetings between the UNICEF Team and OPM provided a forum for discussing the evaluation approach and draft outputs and reporting on progress and challenges. Within the OPM team, the team leader and project manager were responsible for delivering the project to the expected quality and to budget, with ultimate accountability for the project sitting with the Data Innovations Hub at OPM.

The OPM team also undertook periodic consultation meetings with key stakeholders to ensure strong engagement, collaboration, and coordination, which translated into improved stakeholder ownership and will hopefully facilitate more effective implementation of recommendations from this study. This includes representatives from UNICEF Kazakhstan, with whom extensive consultations were held through the life cycle of the project, especially at the inception phase. Also, feedback was sought from the officials at the Government of Kazakhstan and the Ministry of Education on the preliminary findings at an interim presentation, based on which the analysis plan for the project was finalised. The OPM team also spoke to three Non-Government Organisations (NGOs) in Kazakhstan with more than five years of experience of working in the social services delivery to children and young people with special educational needs.[6]

The evaluation was subject to a three-layered quality assurance (QA) process to ensure the robustness and integrity of our results:

[6] To promote active participation, the participants were promised that their identify would not be revealed and have hence been anonymised.

1. A systematic human validation process was done to double-check that social media data scraped from the internet was relevant, in line with Principle 4 of 'validate, validate, validate' (Grimmer & Stewart, 2013).

2. All research products went through a round of internal peer review (i.e. within OPM), conducted by a senior evaluation expert.

3. A presentation on interim findings to UNICEF and the government counterparts was made on 12 July 2023. At the end of the project, research products will finally be reviewed by stakeholders, which in this case means UNICEF (via UNICEF's QA process) and government counterparts, if possible. Comments from key stakeholders will provide an external 'sense check' on the findings and the relevance of recommendations made.

2 Evaluation objectives, framework, and EQs

2.1 Purpose and objectives

The general purpose of the study is to generate insights which can be used by the Government of Kazakhstan to identify existing gaps in the inclusive education landscape and provide evidence to inform future strategies and work in this area. The findings from the study are expected to feed into the next education policy framework and contribute to the implementation of the national project entitled 'Educated Nation' 2021–25. This will be achieved by answering EQs that were finalised at the inception phase and listed in Section 2.1.

This research contributes to identifying the existing gaps and challenges faced by children with SEN, which require improvement. This information will help foster collaboration, engage stakeholders in the policymaking process, address their concerns, and effectively allocate resources to develop targeted initiatives that address the identified gaps. All these elements will feed into the policy cycle, which is discussed in Kazakhstan's context.

More specifically, the primary purpose of this research is to use SML and quantitative text analysis tools to lend a voice to, and analyse the perspectives of, individuals and organisations—in particular parents, teachers, youth, and the general online public—on the status of inclusive education in Kazakhstan. We postulate that, by quantitatively analysing non-traditional, online news and social media sources of data, we will be able to gain insights and formulate recommendations based on viewpoints from a number of different stakeholders that are otherwise difficult to observe or neglected when studied through traditional methods and approaches.

2.2 Policymaking cycle in Kazakhstan

The policymaking cycle in the field of inclusive education in Kazakhstan involves several key stages.

It starts with the initiation of new legislation or a normative legal act specifically focusing on inclusive education. A working group consisting of representatives from the Ministry of Education, education experts, and other stakeholders is formed to develop the new legislation or normative legal act. They conduct research, gather data, and analyse existing practices to inform the development process.

Once the new legislation or normative legal act is formulated, it goes through stages of coordination and expert review. This includes seeking approval and support from relevant government bodies, such as public councils, local executive bodies, the Ministry of Education, and other ministries, as well as parliament. The new legislation or normative legal act is also published on open legal platforms for public discussion and feedback. Amendments or consultations with various stakeholders may be necessary before the new legislation or normative legal act is officially adopted.

After adoption, the new legislation or normative legal act moves into the implementation stage. This involves translating the provisions into specific steps and allocating resources,

including funding and personnel, for effective implementation. Mechanisms for monitoring and evaluation are established to assess the effectiveness and outcomes of the new legislation or normative legal act. This stage includes data collection, research, and the use of relevant indicators to evaluate the progress made in achieving the goals of the new legislation or normative legal act.

Based on the findings from the monitoring and evaluation stage, periodic reviews of the new legislation or normative legal act are conducted. These reviews allow for adjustments, improvements, and revisions to the policy. Feedback from stakeholders, experts, and interested parties is taken into account to ensure that the new legislation or normative legal act continues to meet the changing needs and challenges of the field of inclusive education.

2.3 Groups of interest in this present study

As noted previously, in parallel to the present evaluation study, a formative evaluation to assess the contribution of different state education programmes in promoting inclusive education is already underway (the Junction Bulgaria study). Our research complements this work and provides evidence on the inclusive education landscape in Kazakhstan by using information from publicly accessible social media and online news media sources with particular focus on two groups of stakeholders:

1. the **general online public**, which includes parents, teachers, and youth overall; and

2. **organisations active in inclusive education activities in Kazakhstan** and which possess institutional knowledge on inclusive education in Kazakhstan.

For both these groups of interest, we analyse text that they post or interact with online on inclusive education, identify how they discuss potential barriers, and check if there are any specific policies or programmes that are more frequently mentioned than others. This study also offers a valuable addition to the growing body of literature that uses Big Data to inform policy debates in Kazakhstan. Such Big Data studies have been previously supported by UNICEF.[7]

This research also presents a methodological contribution by synthesising the emerging evidence on the use of innovative analytical methods, such as use of text and sentiment analytics in evaluations. Importantly, in addition to answering specific EQs (listed in next subsection), this study intends to complement the ongoing formative evaluation being undertaken by the Junction Bulgaria team, which can assist the Government of Kazakhstan in identifying existing gaps in the inclusive education landscape in Kazakhstan. This is in tandem with the SDG 4 goal of ‘Ensuring inclusive and equitable quality education for all’. These insights are expected to be useful for the Government of Kazakhstan, which has previously committed to ensuring access to inclusive education and has ratified international agreements that safeguard the rights of children with disabilities, including the right to receive quality inclusive education. For example, in 2015, Kazakhstan ratified the Convention on the Rights of Persons with Disabilities, whose mission is to create conditions where individuals with disabilities can participate in society on an equal footing with others, without any discrimination. Moreover, the Government of Kazakhstan has developed and implemented measures as part of the National Plan to uphold the rights and improve the

[7] Earlier studies have, for example, studied connectivity disparity across schools using Big Data, and were also supported in this endeavour by UNICEF: Unicef Giga Kazakhstan (accessed 11 July 2023).

quality of life for individuals (including children) with disabilities by 2025. This plan represented a significant step towards providing equal opportunities and inclusion.

2.4 Evaluation scope

In terms of the scope of this evaluation, this study focuses on children with development disabilities and relies on a comprehensive definition and conceptualisation of inclusive education (Section 1.1). As the social media footprint and the nature of discussions are likely to vary across groups, we study the perspectives of two vital stakeholders: the general online audience, such as teachers, parents, and students, and specific organisations working in the space of inclusive education in Kazakhstan. Here, we focus on inclusive education at the pre-primary, primary, and secondary level.

The evaluation does not cover an assessment of state programmes as this work is already being catered to in the parallel formative evaluation. Wherever possible, we present geographic insights, but we do not undertake disaggregated analyses across geographies. This is because our experience tells us that geolocations are not always available in the social media data and are often masked by users due to a variety of considerations, such as privacy. As a result, social media data can often not be disaggregated meaningfully to provide heterogeneity analysis by geographic units.

Importantly, these results are restricted to represent the views and perspectives of those who participate on social media and online news media and are therefore not representative of the general population at the national level. For example, if smartphones are the predominant medium through which people access social media, ownership of such a device becomes an important criterion influencing inclusion or exclusion from this study. If these devices are unaffordable to a section of society, some perspectives will not find a place in this research, and this needs to be complemented with other conventional studies.

Some other important features of the scope of this research are as follows.

1. As explained subsequently, we limit our study to the relatively recent past, from the beginning of 2022 to May 2023. This is primarily because multiple social media platforms provide recent data ranging from one to two years.

2. The scope of our study is also limited by the features of how we collect data; in other words, we only ‘see’ conversations and posts online that are captured by our scrapers.

3. In our search queries, we limit the geography to Kazakhstan in the evaluation framework and EQs by including a ‘location’ query that specifies our results must mention locations within Kazakhstan. This includes discussions about inclusive education in Kazakhstan, both from current residents in the country and from the Kazakh diaspora living outside Kazakhstan’s borders.

2.5 EQs

Based on detailed discussion with UNICEF during the inception phase, this study focuses on three EQs.[8] Each is tackled for the two stakeholder groups listed above, the general online public and specific organisations. The EQs we look at are as follows.

1. **What is the content of social and online media discussions on inclusive education in Kazakhstan?** We interpret this to be a broad descriptive question that tries to understand what social and online media text relating to inclusive education looks like, i.e. what is mentioned online and where (e.g. the most-mentioned words, themes, sentiments, etc. in related text, and who is posting about this). We exclusively and purposefully use quantitative text analysis techniques to describe these text corpora.

2. **What do people mention as factors 'holding back' (barriers) inclusive education?** We interpret this as relating to social and online media discussions regarding issues that affect access to inclusive education negatively in Kazakhstan. Again, we interpret this as a descriptive question about online text related to these issues. We interpret 'barriers' broadly, i.e. relating to factors that affects access to education and rights to education for children with disabilities.

3. **Do people mention state programmes at all?** We interpret this to relate to the question of whether specific spate programmes are mentioned in online conversations, news articles, social media posts, etc.

We use a variety of quantitative analytical tools and descriptive analysis methods to answer these EQs. The detailed discussion on how this analysis was undertaken is provided in Section 3 on methodology. Broadly, to capture and study discussions and posts from different stakeholders, we scraped and analysed data separately for the general online public and for specific organisations working in the inclusive education space. Corpus 1 was collated using automated scraping offered by Meltwater.[9] For Corpus 2, we created self-programmed scrapers that manually pulled discussions on inclusive education from publicly accessible social media platforms.

Table 1 provides a snapshot of methods used to answer each EQ for both text corpora. Overall, we use a range of descriptive analyses ranging from analysing trends in mentions, identifying most used words, algorithm-based identification of positive and negative news stories for Kazakh and Russian, and sentiment and hashtag analysis to inform our analysis. Each of these is discussed in detail in Sections 3.2.1 and 3.2.2 respectively for each text corpus.

Importantly, we acknowledge that some conversations online might be gendered. For each of the text corpora, we therefore try to undertake an analysis taking this into account. We try to verify if descriptive analyses vary depending on which gender one pays attention to. It is important to highlight that this does not correspond to a comprehensive analysis of how inclusive education access or barriers vary by gender – but does give us a sense of how different genders are referred to in the context of inclusive education. We are constrained to

[8] It should be noted that these evaluation questions diverge slightly from the scope described in the Terms of Reference to this technical report (see Annex A). This scope was re-negotiated and limited purposefully, together with UNICEF, during the inception phase of this evaluation.
[9] Meltwater is an SML platform, available at https://www.meltwater.com/en (accessed 10 July 2023).

perform more comprehensive analyses by the fact that identifying the gender of authors of online text is (mostly) not possible. (See caveat 8 in section 4.4 on this.)

Although we acknowledge DAC criteria for evaluations,[10] it is important to emphasise that we do not explicitly use them to structure this study. In this study, we present descriptive findings of online discussions on topics that are relevant for inclusive education programming. Thus, this research represents a formative study, which provides evidence on opinions of different stakeholders—as expressed online—on the inclusive education landscape in the country, which can aid the relevance of future programmes. Simultaneously, we acknowledge that the findings of this report only represent the views of those participants who posted their perceptions on publicly available social media platforms, and thus may not be generalisable to the entire Kazakh population. The results presented in this report hence need to be understood within this narrow scope of our study.

Table 1: Analytical tools used for each EQ and stakeholder group

EQs	(a) Corpus 1 (general): Russian and Kazakh	(b) Corpus 2 (specific)
EQ1: Content of online discussions on inclusive education	*Trends in mentions, top publications by reach in Russian and Kazakh, main sources of such information, top negative and positive news, sentiment and hashtag analysis, list of top keywords*	*Spatial concentration of messages on inclusive education, mentions trends, most commonly discussed words, topic modelling, topic prevalence over time*
EQ2: Barriers to inclusive education		*Topic modelling, most cited words*
EQ 3: Mentions of any specific programmes	N/A	*Trends in mentions of specific programmes*

Note: For EQ3, we did not find any mentions of specific programmes using the automated scraping for Corpus 1. Thus, this analysis was undertaken only for Corpus 2, which was manually created using self-programmed scrapers.

[10] https://www.oecd.org/dac/evaluation/daccriteriaforevaluatingdevelopmentassistance.htm (accessed 01 July 2023).

3 Methods

3.1 Scraping and processing text data from the internet

To answer our EQs from Section 2.5, we relied on SML and quantitative text analysis tools. We applied two approaches, depending on the stakeholder group whose online discussions we wanted to capture, resulting in two text corpora (i.e. two different collections of text data).

i. **Corpus 1** was created using Meltwater, an online platform that offers scraping and analytics of social media and online news data. This collection of text data was intended to capture the **discussions of the general online public** on inclusive education in Kazakhstan. We explain how we scrape—i.e. 'collect'—text data using Meltwater in Section 3.1.1.

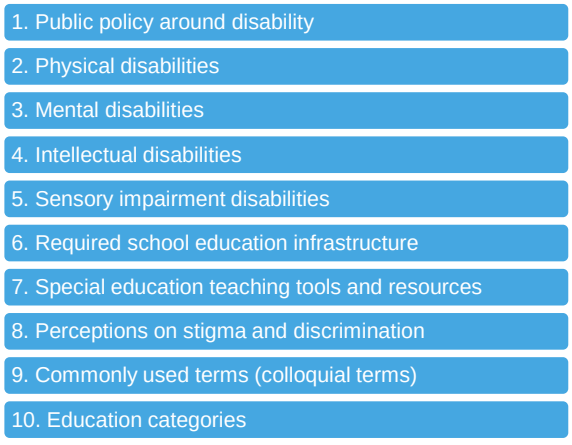
ii. **Corpus 2** was built using self-programmed scrapers and pulled only media **discussions of specific organisations** active in the inclusive education space in Kazakhstan. We explain how we gained access to these data in Section 3.1.2.

Through these complementary approaches analysing similar EQs from different perspectives, we triangulated insights and contrasted these findings across different stakeholders.

3.1.1 Corpus 1: Using Meltwater

As a first step in building a search query in Meltwater, we identified multiple subtopics within inclusive education through a literature review and expert discussions (Figure 2). For each subtopic, we created a combination of keywords in Russian and Kazakh that accurately described the chosen topic and considered the specifics of each language and different stakeholder groups. When compiling our keywords, we took into account declination and conjugation in Russian and Kazakh to encompass all possible variations of words frequently used in media dialogues. An example of keywords for Subtopic 2, 'Physical disabilities', is shown in Box 2.

Figure 2: Subtopics within inclusive education



Box 2: Keywords for Subtopic 2: Physical disabilities

Wheelchair user, muscular dystrophy, spina bifida, cerebral palsy, diseases of the musculoskeletal system, lameness, dwarfism, epilepsy, poor physical coordination, chronic pain, physical handicap, diseases of the central or peripheral neurological systems.

Once our subtopics and keywords were defined, we initiated search queries within the Meltwater platform, scraping online discussions on inclusive education in Kazakhstan. We used the 'combined search' option, which allows us to combine multiple categories of search queries in Meltwater using Boolean operators, i.e. words and symbols such as 'AND' and 'OR', that expand or narrow the search query according to specific parameters. For instance, if we were interested in running a search that includes both Categories 1 and 2 from Figure 2, we combined two sets of keywords using the 'AND' operator. The specific keywords, as well as the use of the Boolean operators in Meltwater, are illustrated in Figure 32 and Figure 33 in the Annex. Meltwater covers the following sources: news sites, blogs, Facebook, Instagram, Twitter, Forums, LinkedIn, Pinterest, Reddit, RSS, TikTok, Twitch, and YouTube. Importantly, in the context of Instagram, Meltwater does not allow extracting the whole text of each post. Thus, this exercise was undertaken as part of Corpus 2, discussed in Section 3.1.2.

As part of the validation process, we initially ran our search queries for a period of 90 days only and assessed the relevance of our results to the topic. To ensure accuracy, we randomly selected a portion of the results and read each of those randomly selected posts or news item thoroughly, classifying them as either 'relevant' or 'not relevant' to the topic of 'inclusive education'. If too many results were considered to be irrelevant, we amended the search query to increase the relevance of results. Multiple iterations of the search queries were conducted to achieve a high percentage of relevant outcomes. Through this process, we were able to increase the number of relevant results in our validation sets from 30% to 77% in the case of searches in Russian, and from 6% to 80% in the case of searches in Kazakh. Once a high level of relevance was attained, the finalised search queries were utilised to collect data over a one-year period, specifically **01 May 2022 to 01 May 2023**.

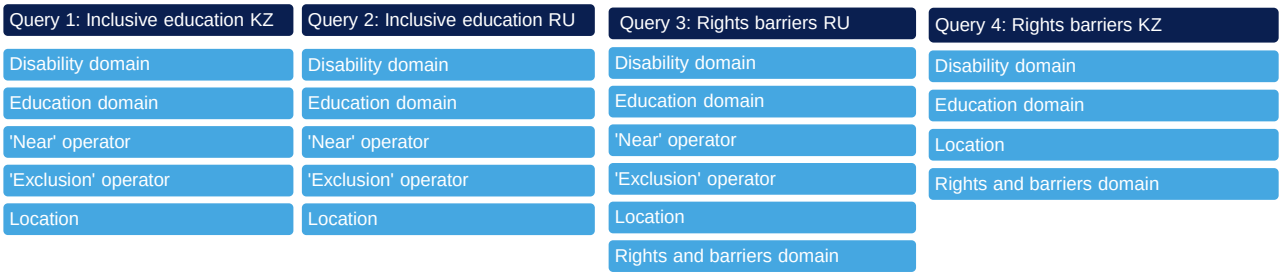
Our final search queries were divided into two categories, both of which were run in Russian and Kazakh. The breakdown of specific subtopics contained within these categories is shown in Figure 3.

1. **Inclusive education:** These combined queries related to subtopics of 'disability' and 'education'. These queries incorporated relevant keywords that specifically covered Topics 2, 3, 4, 5, 6, and 10 from the list of subtopics in Figure 2. To limit our analysis to results in Kazakhstan, we also included a 'location' query that specified our results must mention locations within Kazakhstan. To enhance the quality and relevance of our data, we implemented the 'exclude' function, specifically designed to filter out spam websites. Finally, the 'NEAR' operator ensured that all relevant keywords were within a specific proximity to each other, for instance no more than 50 words apart.

2. **Rights barriers:** This contained the subtopics of 'disability' and 'education' within Kazakhstan, but additionally included queries specifically focusing on Topic 8, 'Perceptions on stigma and discrimination', from Figure 2, defined as the 'rights and barriers domain' search in Figure 3. This category covered keywords relating to stigma and discrimination, as well as violence and violation of rights, to gain insights

into potential challenges and barriers faced by children with spatial educational needs. The slight differences in how our 'rights and barriers' query was defined in Russian and Kazakh emerged from the validation process, as in each iteration we selected the query that gave the highest proportion of relevant results. Importantly, as this category was built within the topic of 'disability and education', some of the themes that featured prominently as topics of discussion may, by construction, refer to barriers in inclusive education. We revisit this point in Section 4.

Figure 3: Key subtopics in our final search queries



Furthermore, we incorporated queries specifically focused on gender, which were not included in the previous topics. These queries consisted of separate searches for gendered keywords associated with girls and boys within Meltwater. However, we discovered a low relevance of the collected publications regarding gender issues, as these publications did not contain specific information or substantial insights related to gender in the context of inclusive education in Kazakhstan.

The number of results in our final search queries for all four categories from Figure 3 are shown in Table 22. Query 2, run in Russian, yielded the largest number of mentions of the topic of inclusive education, with a total of 1,200 mentions. Query 1, run in Kazakh, was mentioned within online discussions 515 times. Given 'rights barriers' is a more specific search intended to pick up discussions focusing on stigma, discrimination, and violation of the rights of children with SEN, we see a lower number of mentions than for inclusive education. As shown by Queries 3 and 4, the topic of rights barriers has more mentions in Russian than in Kazakh, with 326 and 258 total mentions respectively.

Table 22: Number of results from key search queries

Search query	Query 1: Inclusive education KZ	Query 2: Inclusive education RU	Query 3: Rights barriers RU	Query 4: Rights barriers KZ
Number of mentions	515	1,200	326	258

3.1.2 Corpus 2: Using self-programmed scrapers

In addition to capturing perspectives of general public, we also looked at specific organisations working on inclusive education in Kazakhstan. By virtue of their experience and institutional knowledge on inclusive education, their perspectives may differ from those of the general audience and are thus an integral part of the study.

Through a purposive research design, we identified multiple organisations and their publicly accessible social media handles and websites (Annex, Table 6), which were scraped via a self-written programme to create Corpus 2. This means we wrote a computer programme using the programming language Python, which accessed the social media sites and websites and downloaded publicly accessible text. A major component of Corpus 2 came from Instagram channels, Telegram groups and channels, and web sources such as blogs and news outlets. Importantly, this 'manual' scraping approach was also beneficial, as the Meltwater approach presented above did not scrape most of the above sources except for Instagram. Manual scraping of Instagram channels also allowed us to extract the whole text of each post, facilitating a more in-depth analysis.

To account for the exclusivity of discussions, we classified each source as either 'dedicated' or 'not dedicated'. Dedicated sources undertook discussions exclusively on disability and education, whereas the second category shared perspectives on other topics as well. As manual scraping allowed for data extraction over longer periods of time, messages posted between **01 January 2022** and **30 April 2023** were included in the analysis, which meant this timeframe was slightly different than for Corpus 1. Moreover, we only selected and scraped channels and sources identified as part of the purposive design. We applied the following scraping processes to each of the data sources for Corpus 2.

Instagram: Here, we used an external tool called Apify (link) to perform data scraping. Instagram is a closed platform, requiring specialised tools to undertake data extraction, which is often difficult. These data were provided in JSON format. We then extracted selected fields, such as the post text itself, the location of the post (when available), the message, and the author. We anonymised any identification information after analysis.

Telegram: We downloaded these data using the Telegram desktop application. The application does not distinguish between a channel or a group, so we applied the same scraping method as for Instagram. We saved the data to the disk in an HTML format, which were then processed to extract each individual message and the message date.

Web sources: Since online sources vary considerably in their structure, we wrote a separate scraper for each of the web sources. For each source, we first extracted all articles, i.e. all text published on the website. As a second step, we extracted the actual article content together with the article date.

It should be noted, again, that all data scraped using the above processes were publicly available text data. This means we did not scrape any private messages or text published in private groups (Section 3.3).

Once we had completed the data scraping, we saved the results in a local file and further processed it. This involved filtering the entire text corpus by date and translating all the text content into Russian.[11] We used the Yandex translator API for this task (link).

To keep the analysis consistent across the two text corpora, we applied the same queries as in the Meltwater search to each record. In particular, this yielded the following filter: a record was considered relevant for the analysis if it came from a dedicated source or included mentions of both disability and education. From a total of about 12,000 records, about

[11] Due to the purposive selection of sources, the text data was in commonly used languages in the region (either Kazakh or Russian). The decision to translate into Russian was taken because further processing steps required a trained Natural Language Processing (NLP) model to exist for the language. Unfortunately, the library we were most familiar with in this context (Spacy: link) did not include a model for the Kazakh language, and a model could not be found in other common libraries.

10,500 were deemed relevant by using this filter. We also identified about 2,300 and 3,400 key searches on Disability and Education, respectively. Importantly, following these steps, we identified 750 search queries on barriers to inclusive education (Table 33). While for Corpus 1, search queries determined the resulting records themselves, in Corpus 2 search queries were used to select a subset of already-scraped documents. Since analysis on Corpus 2 was only performed in the Russian language, there was no need to specify the language: if the original document was not in Russian, we translated it first and then applied the query.

Next, we applied the NLP model to remove stop words and lemmatise words in the text. This was necessary to perform the next steps of data processing. Wherever available, geographical locations mentioned were extracted from the text.

Table 33: Number of results from key search queries

Search query	Total records scraped from 1st Jan 2022 to 30 April 2023	Dedicated source	Disability	Education	Barriers
Number of Mentions	11891	10181	2277	3424	753

3.2 Data analysis

3.2.1 Corpus 1

As the text data from Corpus 1 was collected using the Meltwater platform, we relied primarily on analytical tools available within Meltwater to conduct data analysis. For both our general search query on inclusive education, as well as the specific query analysing the rights and barriers of children with SEN (Figure 3), we used the following methods.

1. **Mentions trend:** This showed how the volume of posts mentioning our topics of interest evolved over time (for instance, monthly) during the study period. This method allowed us to identify spikes of mentions at a time when online discussion on inclusive education in Kazakhstan was especially active, or (conversely) quiet, and associate them with particular events.

2. **Top publications by reach:** Reach estimates the potential viewership of any particular article based on the number of monthly unique visitors to the specific source. This feature allowed us to analyse which publishing sites had the largest number of viewers, giving us insights into where the most widely viewed discussions on inclusive education were hosted, as well as the ownership of this source.

3. **Top news:** This allowed us to see the content of posts that had the largest reach, and thus to identify the most widely read online publications relating to inclusive education.

4. **Top sources:** This feature displayed the most popular sources for our search queries and included all platforms covered by Meltwater, i.e. news sites, blogs,

Facebook, Instagram, Twitter, Forums, LinkedIn, Pinterest, Reddit, RSS, TikTok, Twitch, and YouTube.

5. **Sentiment analysis:** Using Meltwater's NLP algorithm, this method allowed us to determine whether publications had a negative, positive, or neutral tone. We could also combine this feature with Method 3, 'Top news', and display top news stories that had been assigned a particular sentiment.

6. **Hashtag analysis:** This method allowed us to view the most commonly used hashtags, as well as the number of posts tagged to each hashtag.

7. **Keywords:** This allowed us to view the most commonly occurring words, together with their frequency. We manually excluded 'stop words', such as 'and' or 'the', which did not add any meaningful context to our results.

8. **Gender analysis:** To check if references to one group (girls or boys) had a disproportionately higher representation over the other in discussions on inclusive education (disability and education), we compared the shares of messages referring to each group across two text repositories of 'education and disability' and 'education' respectively. The latter category did not include 'disability' and thus served as a reference or baseline category. For example, if shares of messages and text referring to girls were lower for 'disability and education' *vis-à-vis* 'education', we could infer that issues relating to the experiences of girls were getting lower representation in discussions on inclusive education (disability and education) when compared to discussions on education with no mention of disability. To check for any differences in sentiments across references to boys and girls, we also conducted a sentiment analysis by gender for both text corpora (Section 4.1.3).

3.2.2 Corpus 2

We used multiple quantitative analytical methods to study the text in Corpus 2 extracted from specific organisations working in the inclusive education space in Kazakhstan. These, alongside their purpose, are listed below.

1. **Number of messages per date as trend:** Similar to the mentions trend above, these facilitated studying trends of discussion on topics of inclusive education and how they varied with the academic session.

2. **Word frequency analysis:** This analysis provided a visual snapshot of 50 most common meaningful words that featured in discussions on inclusive education.

3. **Topic modelling using Latent Dirichlet Allocation:** Topic modelling is a statistical exercise that allows us to identify cluster of similar words within a body of text. This analysis cannot be executed on Meltwater and so was exclusively undertaken for Corpus 2. It allowed us to identify the underlying topics in the corpora of documents and provided a way to represent each document as a distribution over the identified topics.

4. **Topic evolution over time:** An extension of the approach above, this method reflected how discussions under each thematic area varied over time.

5. **Messages per country or region normalised by population:** In cases where geolocations could be identified, this analysis provided a geospatial view of the engagement on these discussions.

6. **State programme mentions:** To check if certain programmes were being discussed more than others on social media, we prepared a comprehensive list of active programmes and checked if they were mentioned in the text corpus. This analysis was specifically used to tackle EQ3 as listed in Section 3.

7. **Gender analysis:** As for Corpus 1, we undertook a gender-based analysis in Corpus 2 (Section 4.2.4).

Importantly, to study barriers impeding inclusive education, we again used the methods stated above but against a different set of data, that was scraped using a modified search query to incorporate barriers to inclusive education (Section 3.2.1, Figure 3).

3.3 Ethics

As a values-driven organisation, OPM is always respectful of the rights of the participants of its research projects, and has a policy to ensure complete adherence to research ethics.

We follow a set of ethical principles in conducting all fieldwork based on our own experience and as per the United Nations Evaluation Group evaluation policy. This study also adheres to international best practice standards in evaluation, and is consistent with the OECD DAC framework, even though we do not evaluate specific programmes. We did not conduct any primary data collection as part of this project, but publicly available online data were scraped from the internet. Ethical considerations for using these online data were taken into account, with a particular focus on privacy protection, while again noting that all text we access online is publicly available. This also means that we anonymise the identities of individuals but keep the posts of organisations working on inclusive education in Kazakhstan, as these were posted through their official public social media handles.

As part of OPM's internal protocols, we considered the core principles outlined by the Government Social Research on the use of social media data (2016). Moreover, as per our discussions with UNICEF, we checked again with the Guidance and Evaluation Specialist from ECA regional office that an ethical clearance would not be required for this project, as we were only using publicly accessible social media and news data, as well as anonymising any personal-level identifying information.

3.4 Human rights, gender, and equity

Consistent with the United Nations Convention on the Rights of the Child, particularly Article 28, and the United Nations Convention on the Rights of Persons with Disabilities, particularly Articles 24, 29, and 30, we adopt a human rights-based approach to this study. Here the focus of our research is on equal opportunities for disabled children in accessing pre-primary, primary, and secondary education institutions.

In terms of the legal framework, there is no separate law on inclusive education in Kazakhstan but there exists a President's Law on amendments and additions to some

legislative acts of the Republic of Kazakhstan that relates to inclusive education.[12] As per the earlier version of Law on Education (2007), children with special educational needs were defined as “persons (children) who experience permanent or temporary difficulties in obtaining education due to health, who need special, general education programs of additional education”. Thus, this definition of children with special educational needs led to understanding that inclusive education was interpreted and identified only with respect to children with disabilities. However, in 2021 the definition of children with special educational needs was changed: “persons (children) with special educational needs – persons (children) who have permanent or temporary needs in special conditions for obtaining an appropriate level of education and additional education”. This corresponds to a broader understanding of inclusive education. This 2021 law has become an important document for providing children with special needs with equal access to quality education. The law regulates the activities of psychological and pedagogical support services for children with special educational needs in general education schools, flexibility of curricula, programs, as well as the system for evaluating student achievements depending on the characteristics of their development.

This report explicitly focuses on barriers to inclusive education through different approaches, trying to identify discussions or text posted online that tackles and refers to aspects impeding equitable access to education (Sections 4.1.2 and 4.2.2 on findings). We also check for gender-based differences in our analysis. We check if one group gets a disproportionately higher voice over the other in discussions on ‘education and disability’ when compared to discussions in the reference category of ‘education’. These aspects are outlined in Sections 4.1.3 and 4.2.4 respectively.

3.5 Stakeholder participation

At the inception phase, we held stakeholder consultations with UNICEF and the Junction Bulgaria team. During these consultations, the EQs and the scope of the analysis were finalised. Throughout the implementation of the study, regular biweekly catchups were held with the UNICEF Kazakhstan Team to share project updates and gain regular feedback. As this work stream was expected to complement the formative evaluation, we interacted with the Junction Bulgaria team to gain insights into the progression of their work, and reviewed their preliminary report to identify the intersections between the two workstreams. These are detailed in Section 4.3. Importantly, preliminary findings from our work were presented to UNICEF and the Ministry of Education at a joint workshop on 12 July 2023. This provided us with an opportunity to get input and comments from different stakeholders on the direction of this research and the preliminary findings.

3.6 Methodological limitations

We apply innovative SML tools as our main methodology for undertaking this evaluation. While the use of Big Data offers an opportunity to include online discussions that would not feature within traditional survey-based evaluations, this approach carries its own methodological limitations. In particular, we note the following.

1. **Instagram and Facebook search queries in Meltwater:** There are limitations to the breadth of search queries for Instagram and Facebook in Meltwater, possibly due to

[12] Закон Республики Казахстан от 26 июня 2021 года № 56-VII «О внесении изменений и дополнений в некоторые законодательные акты Республики Казахстан по вопросам инклюзивного образования» (zakon.kz) (Last accessed March 28, 2023).

the data privacy policies of these social media platforms. In the context of Instagram, we were only able to view the posts of business accounts that we specifically tracked. These could be added by manually selecting up to 30 hashtags and up to 30 accounts. For Facebook, we were able to view all public discussions, including comments. However, if we wanted to track the activity of specific pages, there was a limit of 100 accounts. Therefore, it is possible that only a very small subset of the discussions on these platforms was picked up by Meltwater. We address this limitation in Corpus 2, where we use self-programmed scrapers to access Instagram posts directly.

2. **Meltwater analytical tools:** While the online platform offered a wide range of analytical capabilities, outlined in Section 3.2.1, it did not allow the user to download the raw text corpus from its online searches or view the post in full, limiting our ability to conduct independent analysis. As a result, the analysis undertaken on Corpus 1 had to be restricted to the tools offered by Meltwater, with no ability to validate the results using our own methods. This was particularly limiting in the context of gender analysis for Corpus 1, where we were restricted to adding keywords associated with boys and girls to our search, as opposed to methods such as dictionary search, which would have been possible had we had access to the raw data corpus.

3. **Geolocation:** Given the location of most posts was unknown or not disclosed by the user due to a variety of considerations, including privacy, we relied on keywords mentioning locations within Kazakhstan to ensure that the online discussion concerned inclusive education *in* Kazakhstan specifically. For the same reason, we were limited in our ability to carry out an analysis of the spatial distribution of online conversations on inclusive education within Kazakhstan. Despite this, we provide a map of locations of all posts that were geotagged for Corpus 2.

4. **Relevance of results:** We relied on a validation process during which a human expert drew a random subsample of results from our search queries and determined what percentage could be considered relevant to the discussion on inclusive education in Kazakhstan. Given that reviewing several hundred posts individually was unfeasible, we drew conclusions on relevance on the basis of this subsample, rather than the full Corpus. An alternative approach might have been to train a machine learning algorithm to identify the concept of 'inclusive education in Kazakhstan' in text, and classify *all* posts as either relevant or non-relevant. However, this method came with its own set of considerations, such as bias and accuracy in the algorithm, as well as requiring a large training sample.

5. **Language translations:** For Corpus 2, we used a pre-trained NLP model to clean the raw text scraped from the web, as well as to perform topic modelling. As the Kazakh language was unavailable in the Python library that we were most familiar with (SpaCy), and as it could not easily be found among other libraries, we relied on the Yandex translator API (link) to translate all Kazakh content into Russian before analysis. As a result, slight differences in meaning, particularly concerning vocabulary that has a specific cultural connotation, might have occurred.

4 Findings

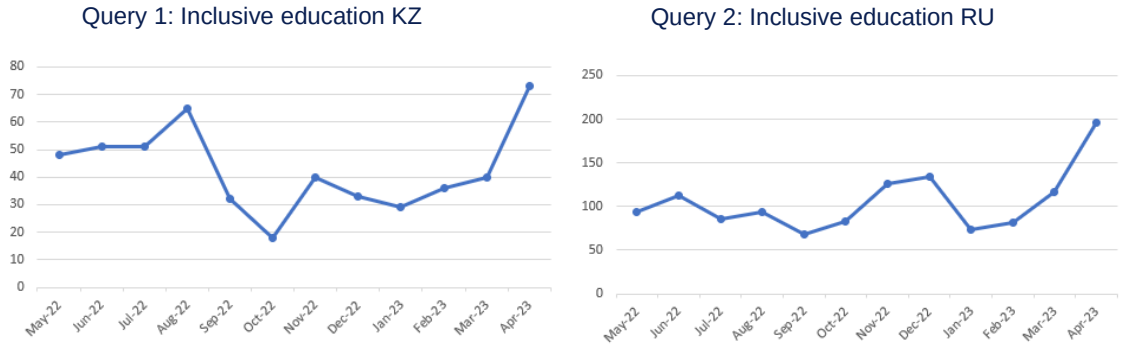
4.1 Corpus 1: discussions by the general public on inclusive education online

4.1.1 What is the content of social and online media discussions on inclusive education in Kazakhstan?

The results from this subsection refer to the first of our two main search queries, Query 1 and Query 2 on inclusive education in Figure 3, and are presented in both Kazakh and Russian. This search encompasses topics relating to disability and education in Kazakhstan and is intended to capture online discussions of the general public on the topic.

We first consider the mentions trend for our two main searches in Russian and Kazakh to see how the volume of online discussion on inclusive education in Kazakhstan changes over time. Figure 4 shows the total number of social media posts that mentioned inclusive education per month for our analysis period (01 May 2022–01 May 2023). The number of mentions varied monthly for both queries, with a dip in mentions in late 2022 and a rise at the beginning of the academic year in September 2023. There was also an increase in April 2023, likely related to the month of raising awareness about autism.

Figure 4: Mentions trends—Queries on inclusive education

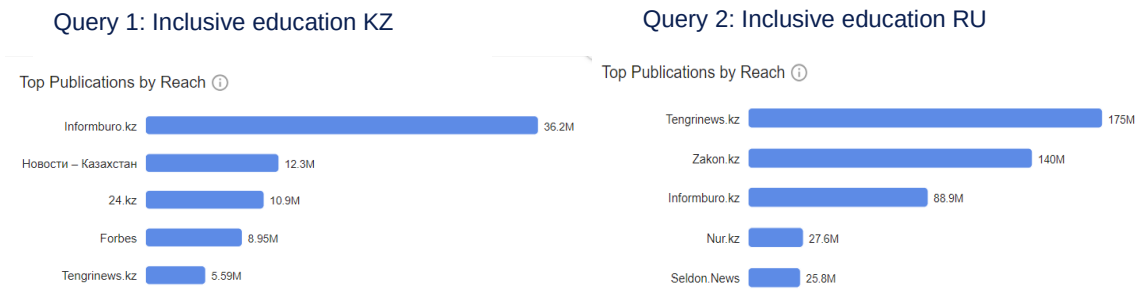


Overall, the number of mentions during our study period was 515 in the case of Query 1 and 1,200 in the case of Query 2. In order to get a sense of the size of the absolute volume of this discussion, we compare these total numbers of mentions to the number of mentions we get from a simple query on only education keywords, covering the same time period. When we run such a query, we get 45,700 mentions for the Russian query and 8,560 mentions for the Kazakh query. **This means that the inclusive education query represents approximately 3% or 6% of the total discussion online on education in Kazakhstan.** Inclusive education therefore does not appear to be a very commonly discussed topic.

Another metric of interest is top publications by reach. According to Meltwater resources, reach estimates the potential viewership of any particular article based on the number of monthly unique visitors to the specific source, indicating which publications likely reached the widest audience. As seen in Figure 5, Informburo.kz could be considered the top source for news on inclusive education in Kazakh. For search queries in Russian, Tengrinews.kz and Informburo.kz appear as the top sources for news on inclusive education.

Both platforms are non-governmental news sources that play a significant role in the media landscape in Kazakhstan, with a substantial audience and influence in the country. This suggests a noteworthy interest in inclusive education among well-known media sources in the region. Our results also show that online posts with the largest number of viewers are likely to be trustworthy and coming from reputable news sources. Both Tengrinews.kz and Informburo.kz have the ability to reach a wide range of readers and generate significant public awareness and engagement around the issues of inclusive education. These news platforms could serve as influential channels for disseminating information and shaping public opinion on inclusive education topics.

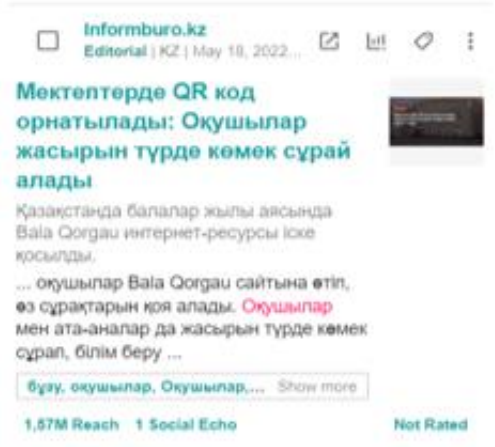
Figure 5: Top publications by reach—Queries on inclusive education



We also look at the top news stories included in the results of our main queries. Top news stories were identified from spikes in mentions trends where the news story with biggest reach was classified as being the top news item. Top news stories in Kazakh (**Figure 6**) mentioned that the Government of Kazakhstan was taking measures to develop inclusive education by introducing modern technologies, creating feedback channels, and protecting the rights of children with SEN. This shows government news has a wide reach, and that the government is successfully highlighting and promoting these actions. This does not necessarily mean these activities are being undertaken at scale by the government. For all the figures in the results section, relevant translations in English are provided as footnotes. According to top news posts in Russian (Figure 7), children with SEN receive attention and support in Kazakhstan, including the provision of housing, improvement of school conditions, the organisation of charity events, and state support through private and public funds.

Figure 6: Top news—Query 1: Inclusive education KZ[13]

[13] Translation of top news titles:
Top row left: Schools install a QR code: students can ask for help anonymously.
Top row right: The winners of the educational grant in 2022 will be announced on August 10.
Bottom row left: Family conflicts, bullying and allowance. The bot Bala Qorgau received more than 500.
Bottom row right: The first EHT in 2023 started on January 10.


Informburo.kz
Editorial | KZ | May 19, 2022...
Мектептерде QR код орнатылады: Оқушылар жасырын түрде көмек сұрай алады
Қазақстанда балалар жылы аясында Bala Qorgau интернет-ресурсы іске қосылды.
... оқушылар Bala Qorgau сайтына өтіп, өз сұрақтарын қоя алады. Оқушылар мен ата-аналар да жасырын түрде көмек сұрап, білім беру ...
бұзу, оқушылар, Оқушылар,... Show more
1,57M Reach 1 Social Echo
Not Rated


Informburo.kz
Editorial | KZ | Aug 9, 2022 · 11:55 AM
2022 жылғы білім гранты иегерлері 10 тамызда белгілі болады
Тізім БАҚ және Ғылым және жоғары білім министрлігінің Telegram арнасында жарияланады.
... толық емес отбасылардан шыққан балалар, сондай-ақ ерекше білім беру қажеттіліктері бар балаларды тәрбиелеп отырған отбасылардан шыққан ...
ерекше білім беру қажеттіліктері
1,3M Reach 2 Social Echo
Neutral


Informburo.kz
Editorial | KZ | Jul 11, 2022 · 10:07 AM
Отбасындағы жанжал, буллинг және жөрдемақы. Bala Qorgau ботына 500-ден астам өтініш түскен
Телеграм-чат арқылы оқушылар арасындағы буллинг фактісі де анықталды.
... зорлық-зомбылыққа ұшыраған баланың анасының өтініші бойынша Алматы облысында тұратын отбасына қамқоршылық органдары, аудан ...
Алматы, бұзу, оқушылар
1,37M Reach 8 Social Echo 1 Duplicates
Not Rated

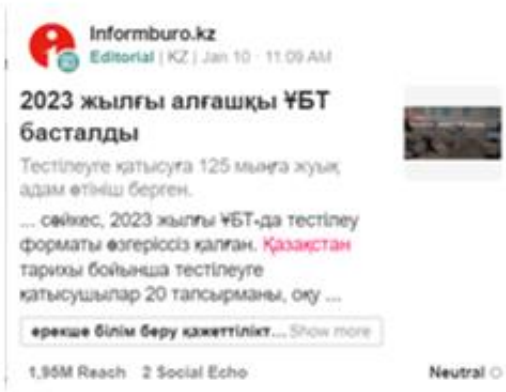
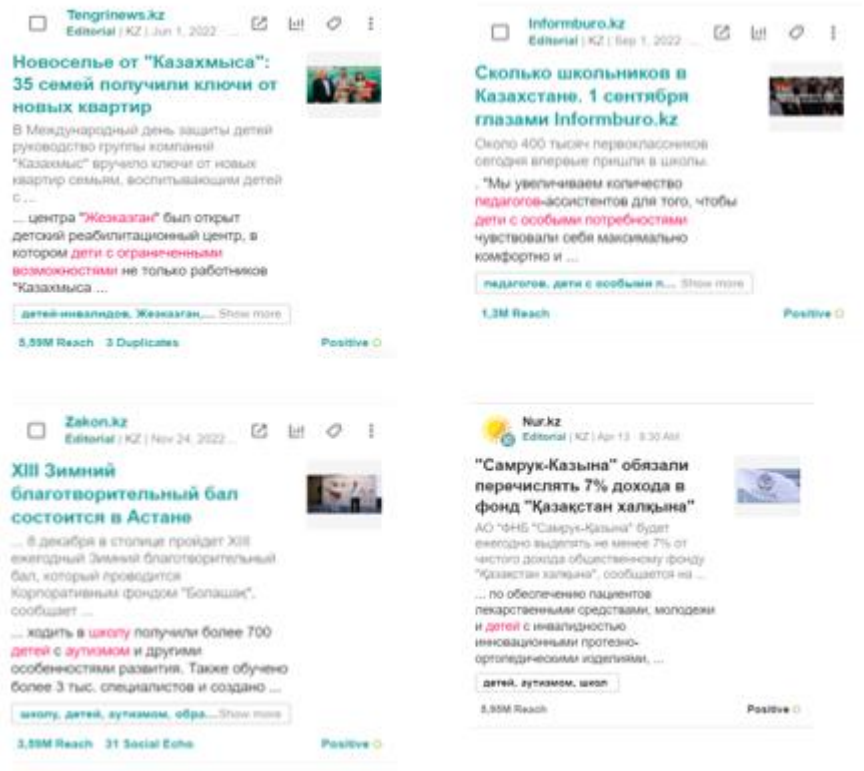

Informburo.kz
Editorial | KZ | Jan 10 · 11:09 AM
2023 жылғы алғашқы ҰБТ басталды
Тестілеуге қатысуға 125 мыңға жуық адам өтініш берген.
... сөйкес, 2023 жылғы ҰБТ-да тестілеу форматы өзгеріссіз қалған. Қазақстан тарихы бойынша тестілеуге қатысушылар 20 тапсырманы, оқу ...
ерекше білім беру қажеттілікт... Show more
1,95M Reach 2 Social Echo
Neutral

Figure 7: Top news—Query 2: Inclusive education RU[14]



An analysis of top sources provides an indication of what types of online sources host the largest volume of discussions on inclusive education in Kazakhstan. As seen in Figure 8, the topic of inclusive education in Kazakhstan is actively covered in news portals and blogs and on Facebook, with very few results drawn from other popular social media sites such as Instagram and Twitter. Overall, the low number of posts coming from social media is perhaps surprising given that civil society organisations working on inclusive education in Kazakhstan primarily use platforms such as Instagram to share their content. However, this result could be explained by the fact that Meltwater can only track specific hashtags and business handles on Instagram that need to be added manually and has a limit of 30 accounts. These limitations are addressed in the analysis of Corpus 2, where we directly scrape Instagram, gaining an insight into discussions among organisations working in the inclusive education space in Kazakhstan.

Figure 8: Top sources—Queries on inclusive education

Query 1: Inclusive education KZ

Query 2: Inclusive education RU

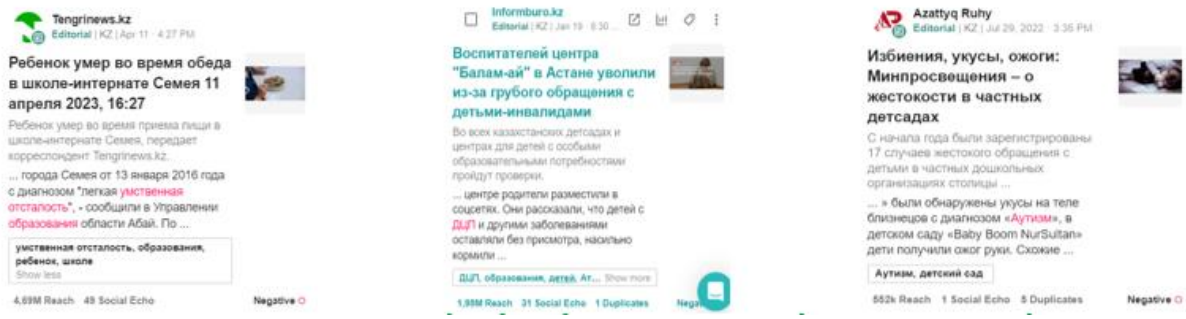
[14] Translation of top news titles:
Top row left: Housewarming from 'Kazakhmys': 35 families have received keys from new apartments.
Top row right: How many school students are in Kazakhstan. September 1 through the eyes of informburo.kz.
Bottom row left: The XIII annual Winter Charity Ball Family will be held in Astana.
Bottom row right: JSC 'National Welfare Fund 'Samruk-Kazyna' was obliged to pay annually at least 7% of net income to the public fund 'Kazakhstan Khalkyna'.

News	422	↗
Blogs	52	
Facebook	25	ⓘ
Twitter	13	↗
Instagram	4	ⓘ
Forums	1	
Comments	0	
LinkedIn	0	

News	1,03k	↗
Facebook	74	ⓘ
Blogs	41	
Twitter	24	↗
Instagram	13	ⓘ
Comments	7	
Forums	5	
LinkedIn	0	

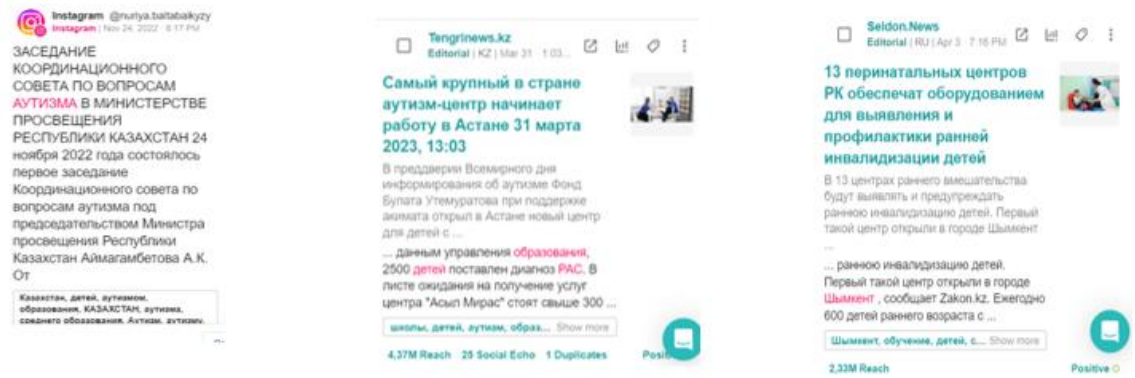
Top news stories with the largest audience reach could also be analysed by sentiment, i.e. the overall tone of mentions determined by Meltwater's NLP algorithm. Sentiment can be either positive, negative, neutral, or not rated if it cannot be adequately classified according to previous categories. Figure 9 shows top news stories in Russian that have the largest audience reach and have been assigned a negative sentiment. Negative news stories reflect discussion on the insufficient protection of the rights of children with SEN and the quality of their education. On the other hand, top positive news stories in Russian, shown in Figure 10, reflect discussion of measures taken to support children with autism and to prevent disability in newborns.

Figure 9: Top negative news with the largest audience reach—Query 2: Inclusive education RU[15]



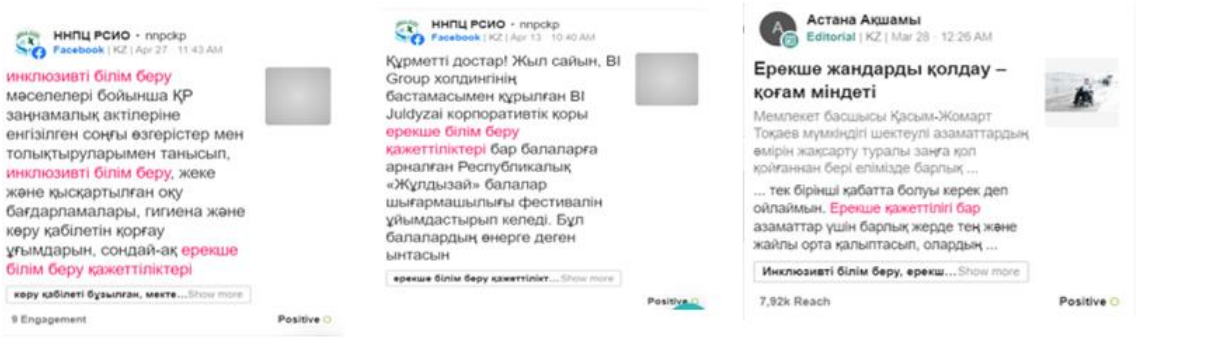
[15] Translation of news titles:
Left: Child died during lunch at the special boarding school.
Middle: The teachers of the Balam-ai centre were fired because of the physical abuse over disabled children in Astana.
Right: Beatings, bites, burns: Ministry of Education—about cruelty in private kindergartens.

Figure 10: Top positive news with the largest audience reach—Query 2: Inclusive education RU[16]



Top positive news stories in Kazakh, shown in Figure 11, reflect discussion on various aspects of inclusive education in Kazakhstan, including the professional development of teachers, various activities aimed at supporting and developing children with SEN, and issues of social protection and support for people with disabilities. Meltwater did not identify any news stories with a negative sentiment in Kazakh.

Figure 11: Top positive news with the largest audience reach—Query 1: Inclusive education KZ[17]



The overall sentiment analysis, shown in Figure 12, tells us the proportion of all posts within the two searches by sentiment. In the case of the search query in Russian, 66.2% of news stories appeared to be positive, while only 3.3% had a negative tone. The remaining 29% of posts were neutral and 1.5% could not be classified and thus were not rated. For results in Kazakh, the majority (73.2%) of posts were neutral, 25.8% were not rated, only 1% had a

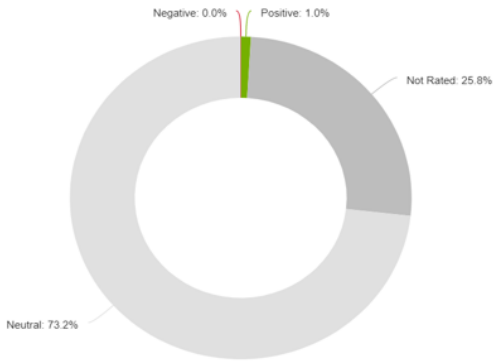
[16] Translation of news titles:
Left: The meeting of the Coordination Council on autism issues at the Ministry of Education of the Republic of Kazakhstan.
Middle: The largest autism centre in the country starts work in Astana on 31 March.
Right: Thirteen perinatal centres of the Republic of Kazakhstan will provide equipment for the detection and prevention of early disability of children.

[17] Translation of news titles:
Left: The latest amendments and additions to the legislative acts of the Republic of Kazakhstan on special and inclusive education considered the concepts of inclusive education, individual and abbreviated curricula, hygiene and vision protection, and what special conditions are created for children with SEN to receive quality education.
Middle: Every year, the BI Juldyzai Corporate Foundation, created on the initiative of the BI Group construction holding, organises the Republican Art Festival of Children, 'Zhuldyzai', for children with SEN. This is a great opportunity for children to show their talent, contributing to increased motivation for creativity.
Right: Support of special people is the duty of society.

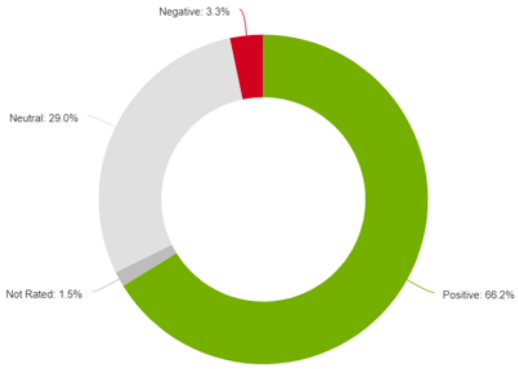
positive tone, and none had a negative tone. This is consistent with there not being any top news stories with a negative sentiment in Kazakh.

Figure 12: Sentiment analysis

Query 1: Inclusive education KZ



Query 2: Inclusive education RU

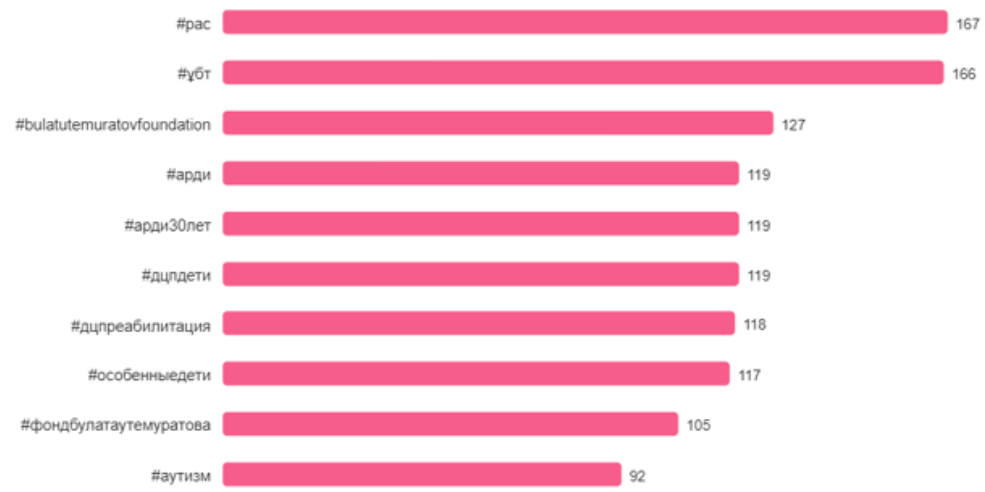


As part of our hashtag analysis, we tracked accounts of specific organisations working in the inclusive education space in Kazakhstan through their Instagram and Facebook handles. The full list of those stakeholders can be viewed in Annex Table 6. According to the top hashtags, shown in Figure 13, autism and cerebral palsy seemed to be the most widely discussed types of disability in the context of inclusive education. Use of the EHT hashtag could demonstrate an active discussion of the issues of adaptation of testing for children with SEN. In terms of the activity of specific organisations, social media users discussed the opportunities and initiatives of the Bulat Utemuratov Foundation, its contribution to inclusive education for children with autism, and the results of its work. For context, some examples of news stories on autism and cerebral palsy are discussed in Box 3.

Box 3: Examples of news items on autism and cerebral palsy

Autism	Cerebral palsy
24 Nov 2022: On 8 December, the capital city will host the 13th annual Winter Charity Ball, held by the Bolashak Corporate Foundation. Dinara Gaplan, Chairman of the Board of Trustees of the Bolashak Foundation, noted that the main purpose of the event is to attract public attention to the problems of obtaining high-quality education for children with autism. **24 Nov 2022:** On 24 November 2022, the first meeting of the Coordinating Council on autism issues was held under the chairmanship of the Minister of Education of the Republic of Kazakhstan, Aitmagambetov A.K. **31 Mar 2023:** The largest autism centre in the country starts work in Astana. On the eve of the World Autism Awareness Day, the Bulat Utemuratov Foundation, with the support of the Akimat, opened a new centre for children with ASD in Astana. This centre has become the second in the capital and the 12th in the Assyl Miras network in Kazakhstan.	**19 Jan 2023:** The teachers of the Balam-ai centre were fired because of the physical abuse of disabled children in Astana. A monitoring group regarding the complaint of parents, consisting of representatives of the Committee for the Protection of Children's Rights of the Ministry of Education, the Territorial Department for Quality Assurance in Education, and the Commissioner for the Protection of Children's Rights, visited the centre of the Charity Fund for the Support of Disabled Children, 'Balam-ai'. Parents complained about the rough treatment by educators of children with cerebral palsy and other disabilities. In order to protect the rights of children, the monitoring group held a conversation with the parents of the affected children and the leadership of the centre about preventing such acts.

Figure 13: Hashtag analysis—Instagram and Facebook[18]

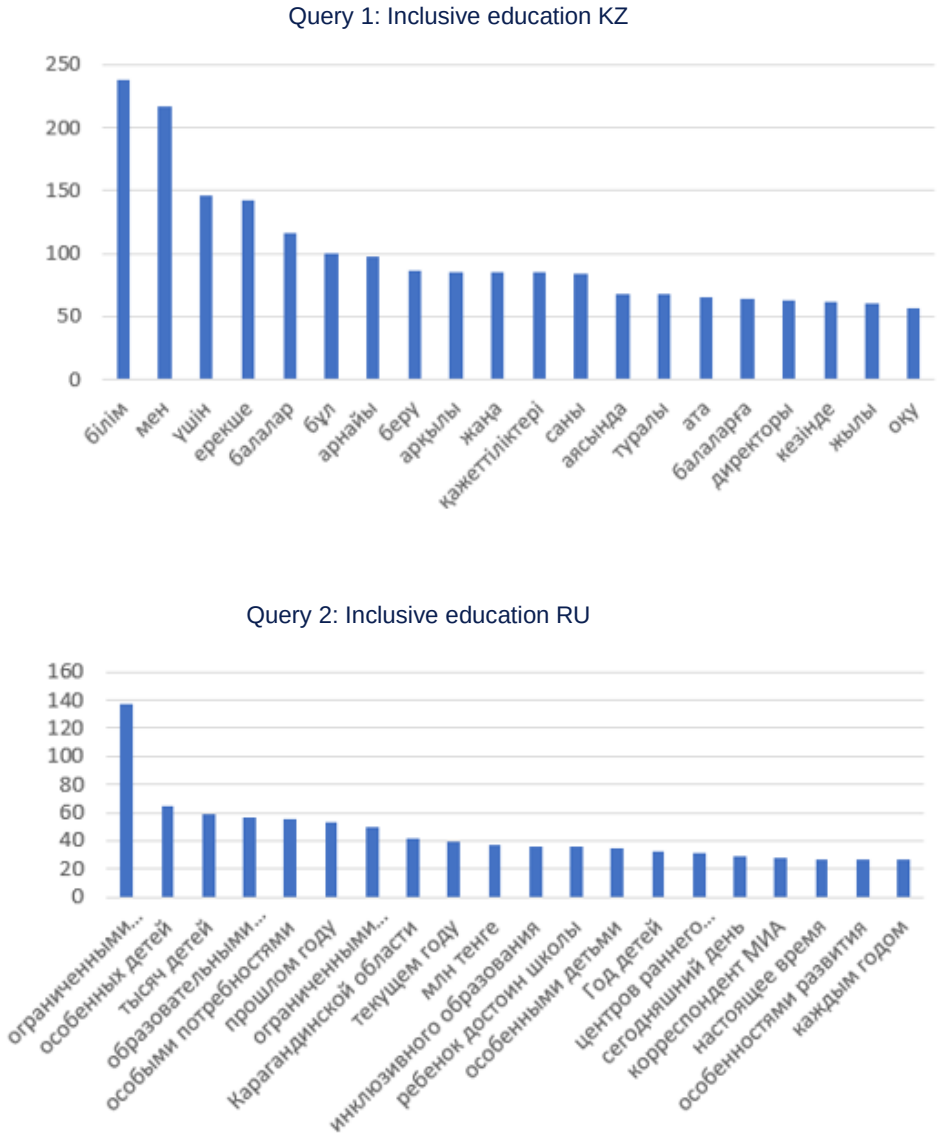


Top keywords, shown in Figure 14, referred to the most commonly occurring words in online discussions on inclusive education in Kazakhstan, together with their frequency. We show the 20 most often mentioned keywords. Popular words reflected discussion on the increase

[18] Translation of hashtags from top down: #asd (autism spectrum disorder); #unt (united national testing); #bulatutemuratovfoundation; #ardi; #ardi30years; #dtspchildren (dtsp: cerebral palsy); #dtspchildren; #specialchildren; # bulatutemuratovfoundation; #autism.

in the number of children with disabilities, measures for preventing disability in children, and mental disabilities.

Figure 14: Top keywords[19]



4.1.2 What do people mention as factors 'holding back' (barriers) inclusive education?

We supplement our analysis of a broader online discussion on topics related to disability and education in Kazakhstan with a specific analysis of Queries 3 and 4, which consider the rights of children with SEN. This subsection will focus on results from search queries that identify particular barriers to inclusive education in Kazakhstan, including discussions around

[19] Translation of top keywords, left to right. Query 1: Knowledge; I; For; Special; Children; This; Special; Through; New; Needs; Number; Within; About; Calling; For children; Director; When; Year; Study. Query 2: Limited; Special children; Thousands of children; Educational; Special needs; Last year; Limited; Karagandy region; Current year; Million tenge; Inclusive education; A child has a right for a school; By special children; Year of children; Center of early; Today; MIA correspondent; Current time; Special developmental needs; Each year.

stigma and violence. To illustrate the content of this search query, Box 4 shows additional keywords included under the 'rights and barriers domain'.

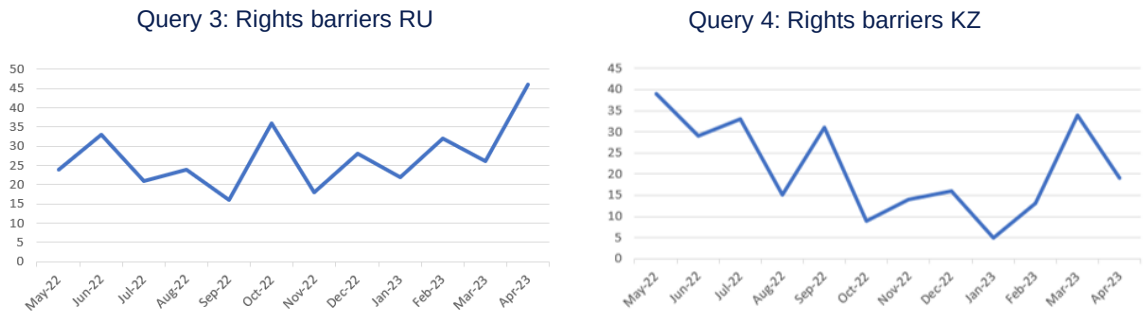
Box 4: Additional keywords associated with the 'rights and barriers domain'

Bullying by parents, teachers, children, social isolation, disconnected from their peers, classmates and society as a whole, discriminatory barriers to higher education, unable to learn, negative stereotypes, contact with other children, lack of culture, interfere with academic performance, isolated from peers, inappropriate, lack of opportunities to continue education, low expectations, harassment, home schooling, negative public attitude, norms, cultural practices (Russian\Kazakh), feeling ashamed or negatively about disadvantages, behaviour around integration and inclusion, life skills, employment for future, earning, agency, society is not inclusive, society is not ready, being segregated, discrimination of children with special needs, to be in society, communication with other children, communication with peers.

The total monthly mentions in Kazakh showed a decreasing trend throughout the year (Figure 15). There were two spikes in September and March coinciding with messages from the President of Kazakhstan mentioning support to people with SEN. The mentions trending in Russian followed a similar pattern (Section 4.1.1), with variation throughout the year and an increasing trend over the academic year, starting in September. Similarly, the increase in April coincided with the month of raising awareness about autism.

Overall, the total number of mentions during our study period was 326 for Query 3 and 258 for Query 4. This is roughly one-quarter of total mentions during the same period for the broader query of inclusive education in the case of Russian, and around one-half the number of mentions in the case of Kazakh. When compared to a broader search concerning education in Kazakhstan, posts relating to rights and barriers accounted for 0.7% of results in Russian and 3% of results in Kazakh. **Therefore, discussions on stigma, discrimination, and violence against children with SEN constitute a very small portion of the online conversation on inclusive education in Kazakhstan, and an even smaller portion of the discussion on education more broadly.**

Figure 15: Mentions trend—Rights barriers



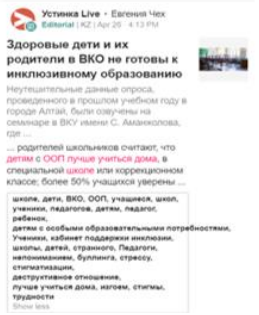
Top news stories from our search on rights and barriers are shown in Figure 16. Top posts in Kazakh mentioned **child abuse** in kindergartens 17 times. Consistent with this theme, another post mentioned that the Government of Kazakhstan had discovered cases of violence and bullying against children and youth at schools and universities, as well as in families, using digital technology (a telegram bot). Because children with disabilities are disproportionately more affected by violence and abuse than their peers without disabilities, their safety at school is critical for adequate inclusion. In response, the Government of Kazakhstan intends to address quality issues, improve standards, and establish minimum safety requirements to ensure the wellbeing and development of children in early childhood

education. This is being done by improving kindergarten standards regarding the qualification of staff members, introducing a licence to run a kindergarten, tightening government contract requirements, and improving the content of early childhood education programmes. These points also feature in the new education concept released by the Ministry of Education.

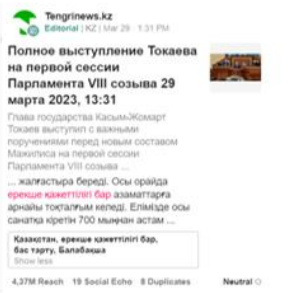
Top posts in Russian stated that the Government of Kazakhstan was putting in effort to provide inclusive opportunities in sports and recreational activities for children with disabilities, as well as to create stricter safeguards and monitoring mechanisms in educational settings to ensure their safety and wellbeing. Another post reported a growing inclusion of children with autism in regular classrooms, highlighting the importance of comprehensive support systems and raising awareness to foster positive attitudes towards inclusive education.

Figure 16: Top news—Rights barriers[20]

Query 3: Rights barriers RU



Query 4: Rights barriers KZ



Top negative news stories in Russian, shown in Figure 17, reflected discussions of the challenges of inclusive education, highlighting incidents of abusive treatment and safety violations in educational settings. On the other hand, top positive news stories in Russian (Figure 18) reflected a positive trajectory in Kazakhstan's inclusive education landscape (accessible school buses, autism centres, and charity awareness events). Autism centres here function with the main objective of identifying ASD and helping children improve their social, speech, and household skills. Honing these skills sets prepares them for being mainstreamed into schools and fosters inclusive education. Meltwater did not

[20] Translation of top news (rights barriers RU):
Top row left: The centre of hippotherapy was launched in Astana on 25 October.
Top row right: Kindergarten teachers who bullied the special child were fired.
Bottom row left: How to stop an unruly and armed teenager.
Bottom row right: Children without disabilities and their parents are not ready for inclusive education.
Translation of top news (rights barriers KZ):
Top row left: Requirements for running kindergartens were strengthened.
Top row right: Family conflicts, bullying, and allowance: Bala Qorgau bot has received more than 500 requests.
Bottom row left: Yerketai kindergarten teachers locked a child in the toilet.
Bottom row right: Full speech of Tokayev at the first session of VIII Parliament.

identify any news stories in Kazakh as either positive or negative. Despite this, the content of top news stories in Kazakh (Figure 19) seemed positive overall. The news stories reflected a positive focus on providing inclusive education resources, promoting sports participation, and improving school accessibility for children with disabilities in Kazakhstan.

Figure 17: Top negative news with the largest audience reach—Query 3: Rights barriers RU[21]

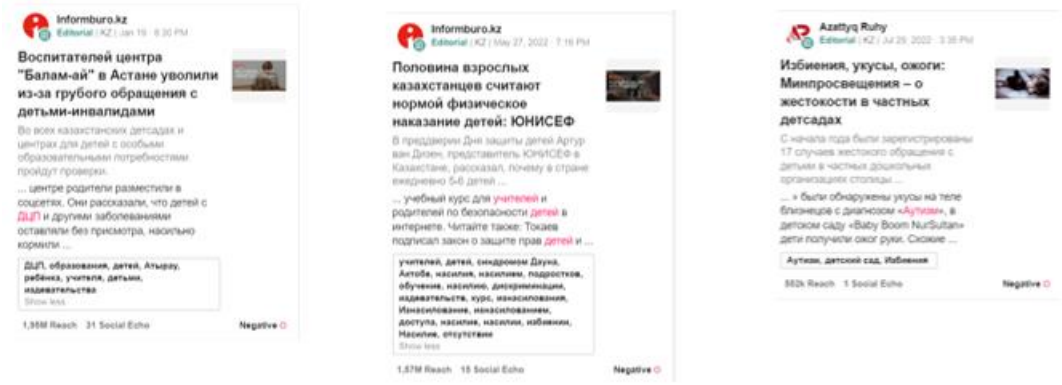
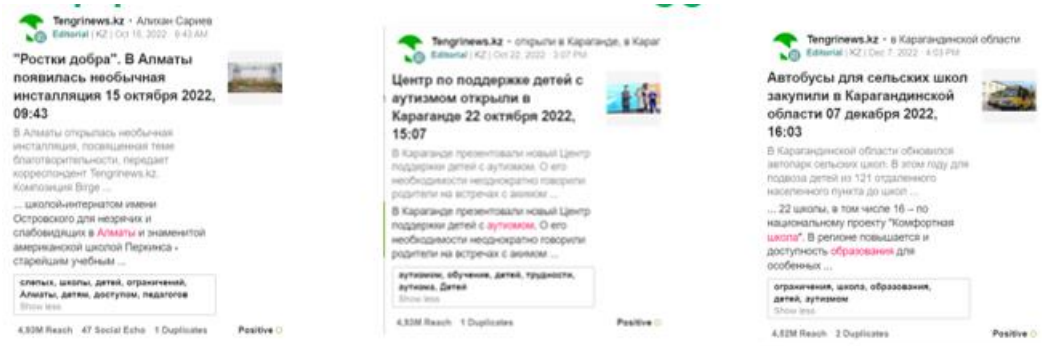
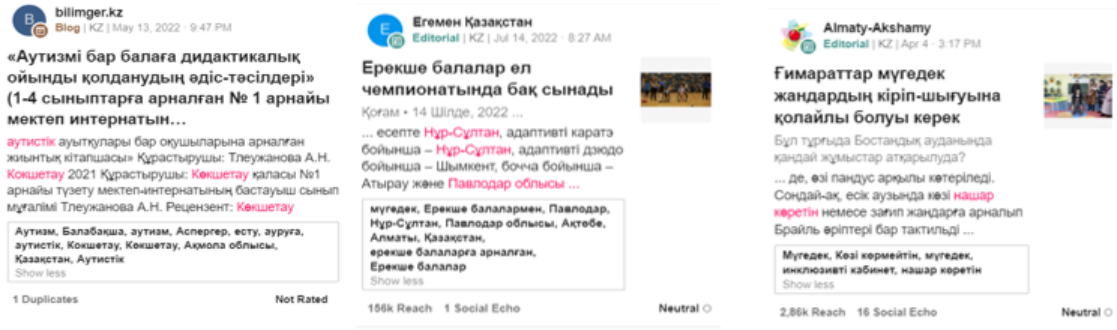


Figure 18: Top positive news with the largest audience reach—Query 3: Rights barriers RU[22]



[21] Translation of news titles:
Left: The teachers of the Balam-ai centre were fired because of the physical abuse of disabled children in Astana.
Middle: Half of Kazakhstani adults consider physical punishment of children to be the norm: UNICEF.
Right: Beatings, bites, burns: Ministry of Education—about cruelty in private kindergartens.
[22] Translation of news titles:
Left: 'Sprouts of Goodness': In Almaty, an unusual installation has been unveiled.
Middle: An Autism Support Centre has been presented in Karaganda.
Right: Buses for rural schools were bought in Karaganda region.

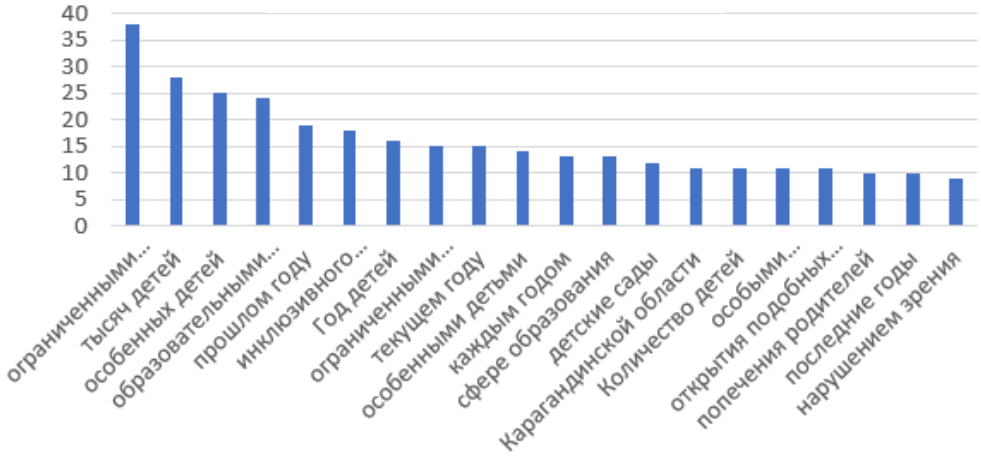
Figure 19: Top news with the largest audience reach—Query 4: Rights barriers KZ[23]



Analysis of top keywords (Figure 20) reveals a growing focus on inclusive education in Kazakhstan over the last few years. Increased attention has been paid to the requirements of children with SEN with the establishment of inclusive schools and efforts to provide care and support. The analysis of popular words in Russian shows that the term '**limited opportunities**' was commonly utilised when referring to children with SEN, despite its offensive nature, even in official documents. As discussions on stigma, discrimination, and violence against children with SEN constituted a very small portion of the online conversation on inclusive education in Kazakhstan, initiatives aimed at the use of inclusive language in official communications and correspondence may help normalise these conversations in the rest of the country.

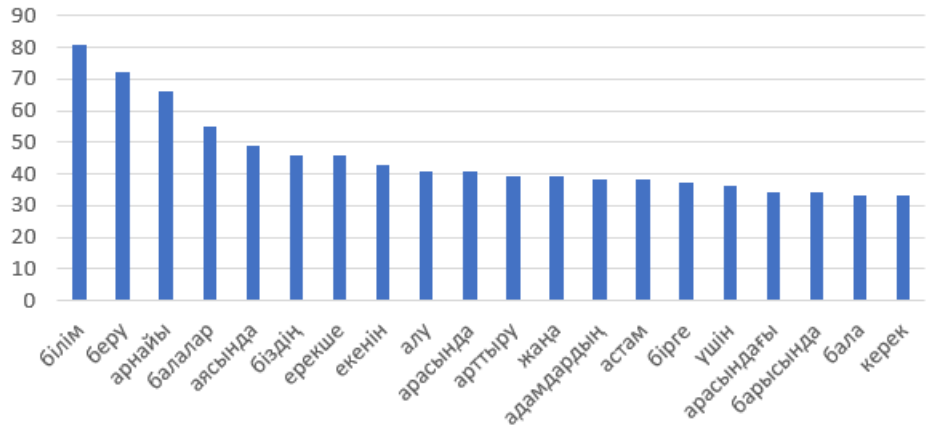
Figure 20: Top keywords (rights barriers)[24]

Query 3: Rights barriers RU



Query 4: Rights barriers KZ

[23] Translation of news titles:
Left: Methods and techniques of using didactic games for children with autism.
Middle: Special children competed in the national championship.
Right: Buildings should be accessible for people with disabilities.
[24] Translation of top keywords (RU), top down:
Limited; Thousands of children; Special children; Educational; Last year; Inclusive; Year of children; Limited; Current year; Special children; Each year; Sphere of education; Kindergartens; Karaganda region; A number of children; Special; Running of similar; Parental care; Last years; Visual disabilities. Translation of top keywords (KZ), top down: Knowledge; Provision; Special; Children; Within; Our; Special; So; Taking; Among; Increase; New; People's ; More than; Together; For; Among; During; Child; Necessary.

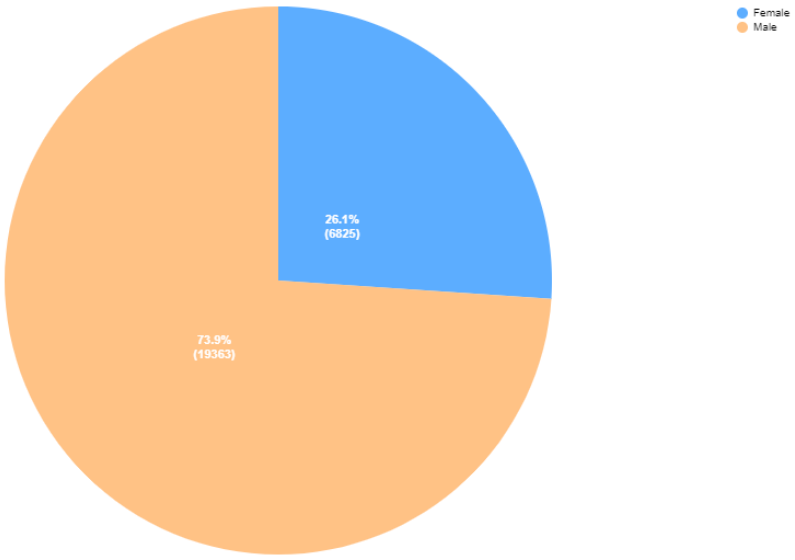


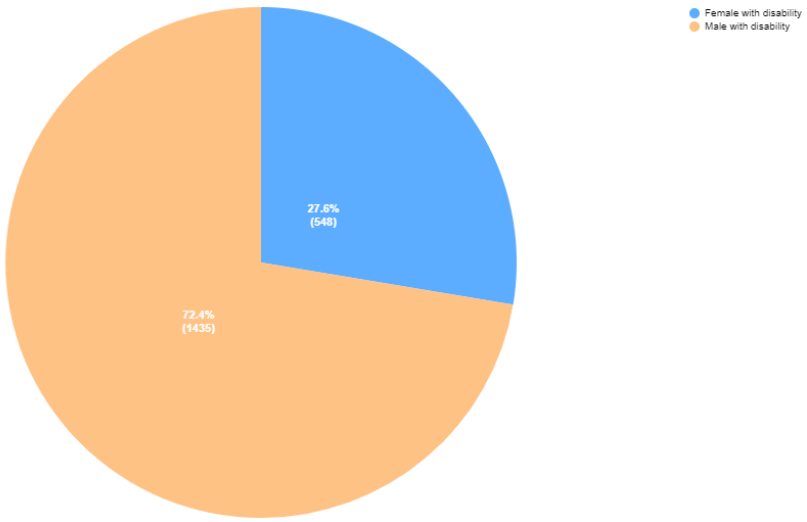
4.1.3 Gender analysis

Existence of gender disparity, which is prominent globally, can cause differentiated access to inclusive education. To check if girls or boys are disproportionately represented in discussions on inclusive education, we compare the shares of messages mentioning each group across the two text categories of inclusive education and education without any mention of disability. For example, if the share of messages regarding girls is lower for 'disability and education' than for 'education', we can infer that girls are getting lower representation in discussions on inclusive education (disability and education) when compared to discussions on education with no mention of disability.

When comparing the share of messages on education and disability with the share of messages on only education for both boys and girls, we find these shares to be comparable, with about 75% of the messages representing boys. This suggests the discussion on inclusive education, which focuses on children with disabilities, does not accord higher importance to one group *vis-à-vis* the reference category. Having said this, it is important to note that, at an absolute level, the volume of messages across time caters to boys (~75%) more than girls (~25%) (Figure 21).

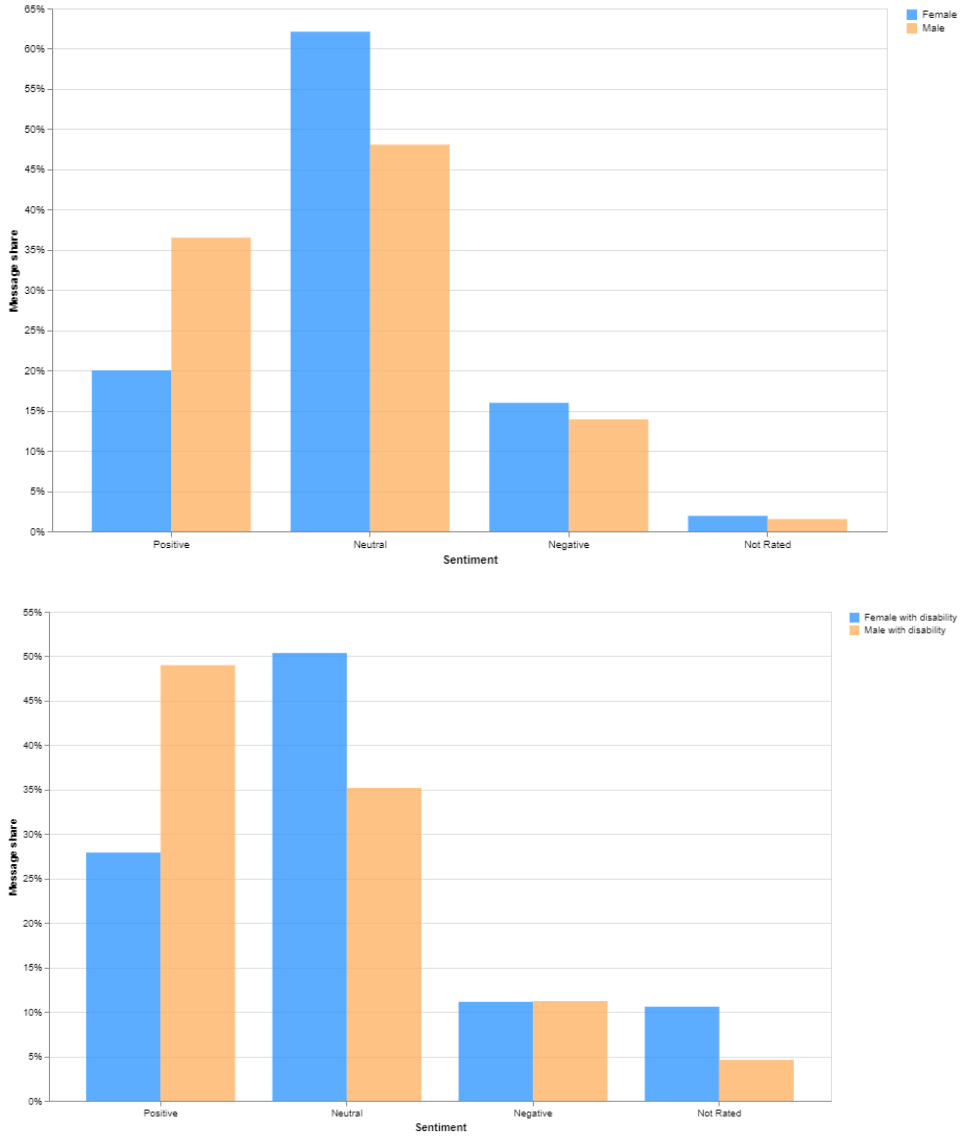
Figure 21: Gender analysis (Corpus 1)





Next, we check for any differences in the sentiment expressed by gender in the education discussion compared to the discussion on education and disability (Figure 22). For both kinds of message, we find that the positive sentiment is expressed more frequently in the context of boys, whereas the neutral sentiment is expressed more frequently in the context of girls. The negative sentiment, which appears to be expressed slightly more commonly for girls *vis-à-vis* boys on education, is expressed less frequently and becomes equal to that of boys for messages on education and disability.

Figure 22: Differences in sentiments across education and education and disability, by gender (Corpus 1)



4.1.4 Summary of key findings

Inclusive education in Kazakhstan is a topic of ongoing discussion. However, it represents only a small portion (3–6%) of the general discussion on education in Kazakhstan, making it a relatively uncommonly discussed topic. Average monthly mentions vary, but there was an overall increasing trend over the academic year 2022–23. The increase we observed in April 2023 coincided with the month of raising awareness about autism in Kazakhstan. Analysis of top news stories shows that the Government of Kazakhstan is actively taking measures to develop inclusive education, including through the introduction of modern technologies, the creation of feedback channels, and initiatives aimed at protecting the rights of children with SEN. This shows that the government's messaging has a wide reach, and that they are successfully highlighting and promoting these activities. In terms of the sources of online discussion, news sites, news portals, and Facebook are actively covering the topic of

inclusive education in Kazakhstan. Discussions on other social media relating to inclusive education were relatively limited.

Negative news stories often focused on the lack of protection of the rights of children with SEN and the quality of their education. On the other hand, positive news stories highlighted measures taken to support children with autism and to prevent disability in newborns, as well as various aspects of inclusive education in Kazakhstan, including capacity building of teachers and measures for social protection. Autism and cerebral palsy were the most widely discussed types of disability in the context of inclusive education.

Hashtag analysis shows that autism and cerebral palsy seemed to be the most widely discussed types of disability in the context of inclusive education. The adaptation of testing for children with SEN was actively discussed, as indicated by the EHT hashtag. Among non-profits, the Bulat Utemuratov Foundation and its initiatives received attention and positive discussion for its contribution to education for children with autism. The most widely used topics in discussions included an increase in the number of children with disabilities potentially linked to better diagnostics, and the equipping of perinatal centres to prevent disability in newborns.

An analysis of 'barriers' reveals that incidents of violence and bullying in kindergartens, schools, universities, and families have been identified by the Government of Kazakhstan, emphasising the need for measures to protect children and youth. Consistent with earlier results, we found that the Government of Kazakhstan has made an effort to address quality issues, improve standards, and establish minimum safety requirements in **early childhood education.** However, in comparison to a broader search on education in Kazakhstan, **conversations on stigma, discrimination, and violence against children with SEN constituted a very small portion of online discussions.** Positive news stories reflected progress in inclusive education, such as improving accessibility of school buses, autism centres, charity events, and efforts to promote sports participation and improve school accessibility for children with disabilities. Autism centres here help children improve their social, speech, and household skills for mainstreaming children into schools. The presence of hashtags relating to charity and support indicates **community involvement** and efforts to enhance the wellbeing of, and opportunities for, children with SEN. Although these charity events can be seen as paternalistic, in the context of Kazakhstan they also play an important role in integrating children with SEN and disabilities into society and are important in changing the existing mindset and how these children are viewed. Inclusive education in Kazakhstan has attracted growing attention, including the establishment of inclusive schools and efforts to address diverse needs, such as visual impairments, within the framework. Finally, the use of the term 'limited opportunities' when referring to children with SEN is common but considered offensive, highlighting the need for more inclusive language in official documents.

When analysed by gender, we find shares of messages for boys and girls across the corpora of education and disability and education to be comparable. However, the share of messages regarding boys in such discussions was about three times as large as that of girls, accounting for about 75% of the overall discussions.

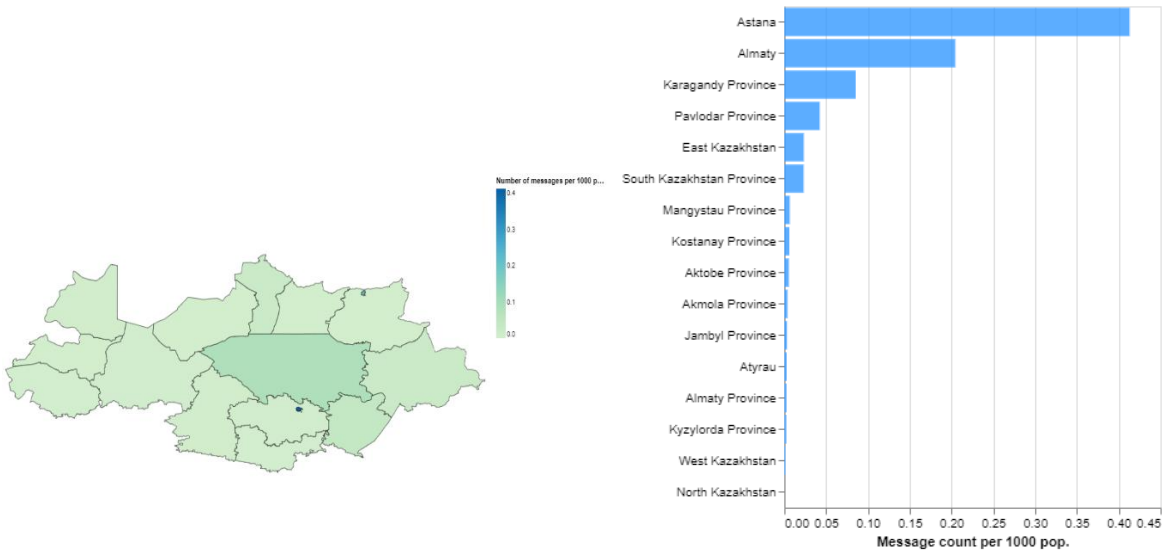
4.2 Corpus 2: Online discussions by organisations focusing on inclusive education in Kazakhstan

4.2.1 What is the content of social and online media discussions on inclusive education in Kazakhstan?

We now discuss the findings from Corpus 2, based on the perspectives and experiences of organisations actively engaged in inclusive education issues in Kazakhstan (see Section 3.1.2 and Table 6 in the Annex for a description of how we built this text corpus and the list of organisations and websites included here). This analysis serves two purposes. First, it allows us to identify commonalities and differences with the online discussions undertaken by the general online audience in Section 4.1. We can thereby compare findings, assess whether the approach to Corpus 1 missed important topics, and identify the themes that are most important to all stakeholders. Second, as these organisations possess institutional knowledge on inclusive education and regularly engage with the Government of Kazakhstan, insights into what these organisations accord importance to is vital to understand how the inclusive education is being discussed and might evolve in the country.

As organisations are more open to geotagging their locations than individuals, we first analyse spatial patterns in online discussions on inclusive education.[25] Figure 23 displays the number of messages per 1,000 population, both in the map (left-hand panel) and the bar chart (right-hand panel).[26] These discussions are found to be concentrated mostly in large cities of Astana and Almaty. Even after accounting for higher population densities in larger cities, predominant social media engagement on inclusive education is higher in these cities.

Figure 23: Spatial concentration of messages on inclusive education



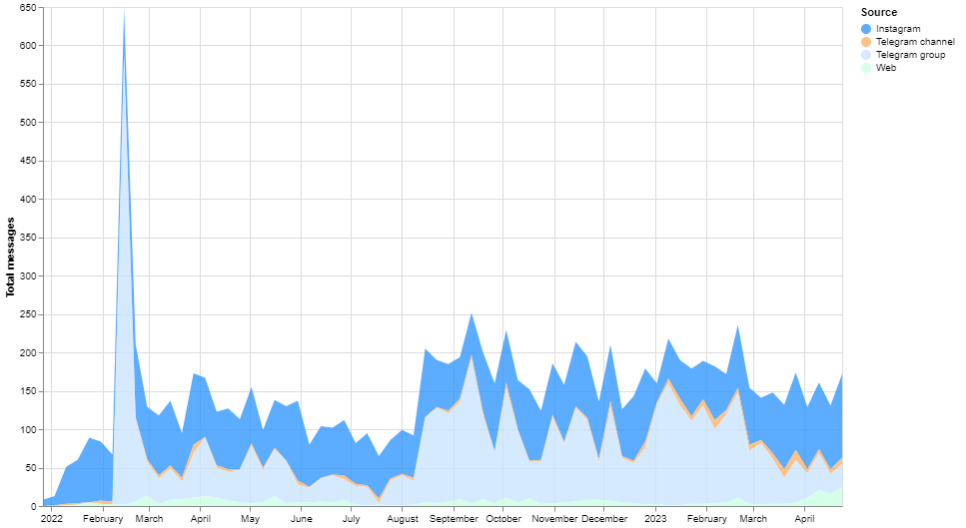
We use the same combined search query on disability and education as we did for Corpus 1 to identify time trends in the volume of online discussion on inclusive education. Consistent with what we found earlier, the number of messages rises as the academic session starts (around September) and tapers off gradually. Since the beginning of 2022, the trend in the volume of discussion appears to be rising. Organisations active in this space mostly use

[25] Only messages that have a geographical reference are included in this analysis.

[26] Population numbers for each region were taken from the Kazakhstan Bureau of National statistics (link).

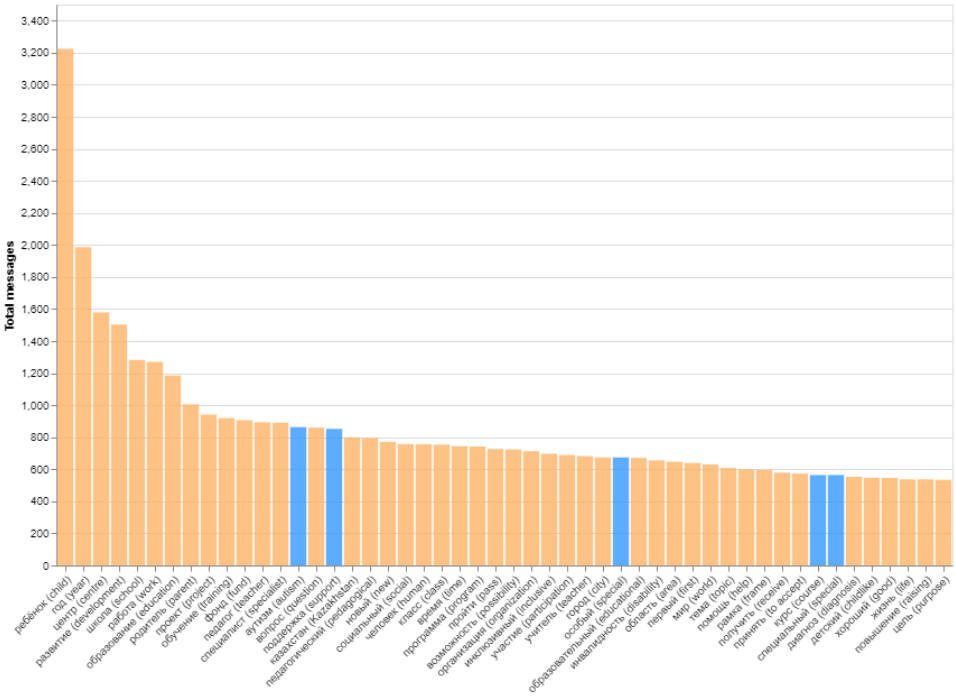
Instagram and Telegram groups for public discussion (Figure 24). Importantly, unlike for Corpus 1, there was no uptick in mentions in April, the month of raising awareness on autism. However, as we show below, there are significant mentions and discussions regarding autism among organisations engaged in inclusive education in Kazakhstan.

Figure 24: Mentions trends for Corpus 2



Next, we analyse the top 50 most frequently used words in the text corpus, not by the frequency of their appearance overall but by the number of documents or messages that contain these words. In this way, we account for varying document lengths and identify common areas of discussion on inclusive education. In terms of types of disability, autism is the most frequently discussed type of disability. Analysis also shows common mentions of terms like ‘special courses’ and ‘support’. This suggests that a big part of the discussion we are observing relates to the capacity building of teachers and specialists and measures to support special education (Figure 25). Next, we use topic modelling algorithm to classify these discussions into a set of common and independent themes.

Figure 25: Most common discussion words



To identify common themes from Corpus 2, we undertake topic modelling, which classifies the most common words into broader topics of discussion. This is done in the following manner:

1. we specify the number of different topics needed to the topic modelling algorithm—for topic modelling analysis in this section, we consider seven most-discussed topics;

2. the algorithm then provides a list of words associated with the seven topics; and

3. based on what words or phrases feature in each topic, we provide a broad label to this category, which is also cross-verified using ChatGPT.[27]

We now analyse the list of 10 most-used keywords for each topic identified in this way. In Table 4, we provide the English translation of the original Russian topics, which are provided in Annex Table 7.[28]

The findings show a focus on the quality of education, with the capacity building of teachers (Topic 1) and school education content and pedagogy (Topic 2) being the most prominent categories. Importance is also accorded to diagnosis, treatment (Topic 6), and social support and opportunities for children with disabilities (topic 7).

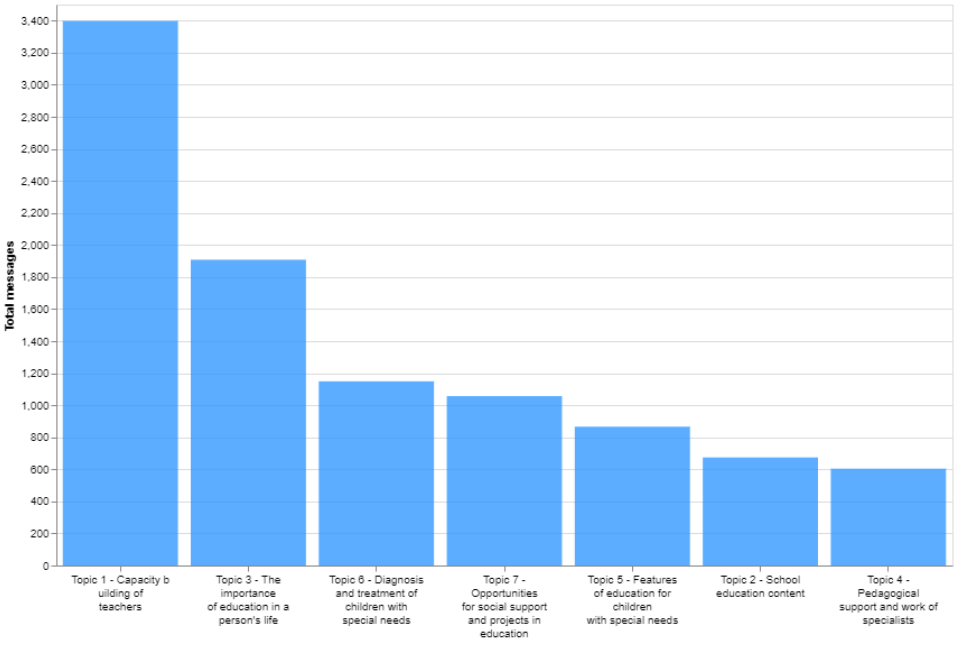
[27] https://openai.com/blog/chatgpt.
[28] Although we were careful with translations, there might have been translation losses from Russian to English. The Russian text is provided in Annex Table 7.

Table 4: Topic modelling (English) for seven topics identified by the algorithm

Keywords in each topic	Capacity building of teachers	School education content	The importance of education in a person's life	Pedagogical support and work of specialists	Features of education for children with SEN	Diagnosis and treatment of children with SEN	Opportunities for social support and projects in education
1	organisation	receive	good	spend	organisation	receive	frame
2	year	class	class	support	pedagogical	pass	possibility
3	development	school	important	result	school	new	world
4	work	development	life	pass	requirement	childlike	city
5	area	topic	time	pedagogical	educational	question	year
6	pedagogical	raising	violation	work	special	child	support
7	education	class	development	year	inclusive	diagnosis	Kazakhstan
8	teacher	training	human	work	special	disability	social
9	school	programme	parent	specialist	child	autism	fund
10	teacher	course	child	centre	education	year	project

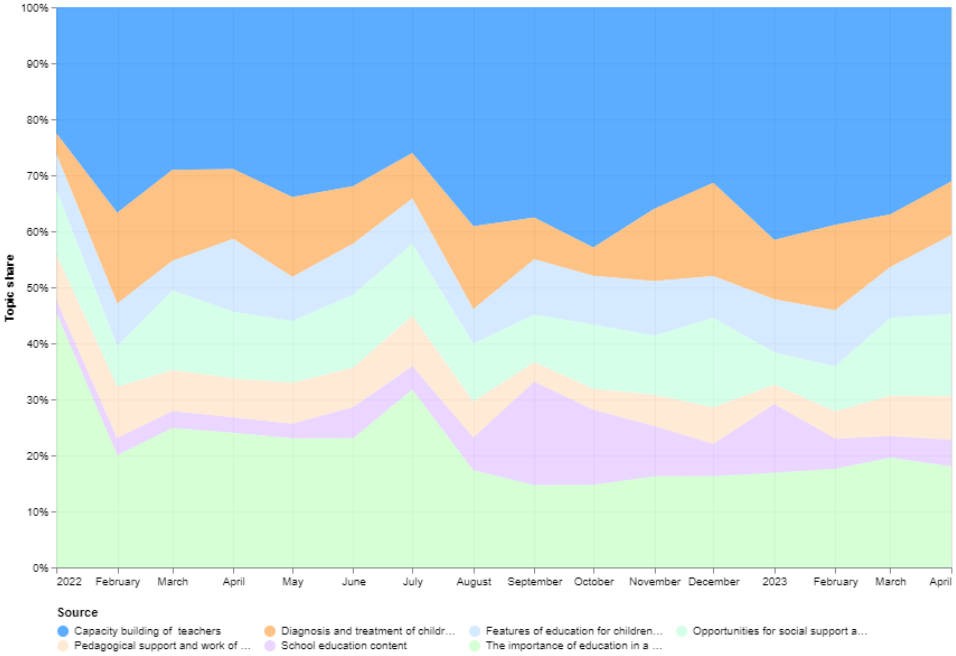
To quantify how often these topics are mentioned in relation to one another, we analyse the number of total messages per topic (Figure 26). We find that the capacity building of teachers is the most important topic, and that the number of messages on this theme (3,400) is close to twice as large as the second most-discussed topic (school education content) (1,900).

Figure 26: Messages per topic



Next, we check how mentions of these topics change over time by analysing the prevalence of topics per month (Figure 27). We find that the importance of certain topics, such as the quality of education and the capacity building of teachers, is sustained over time and that there are no strong seasonal patterns. However, the time series length may be too short for this analysis, and patterns may become apparent when using longer intervals.

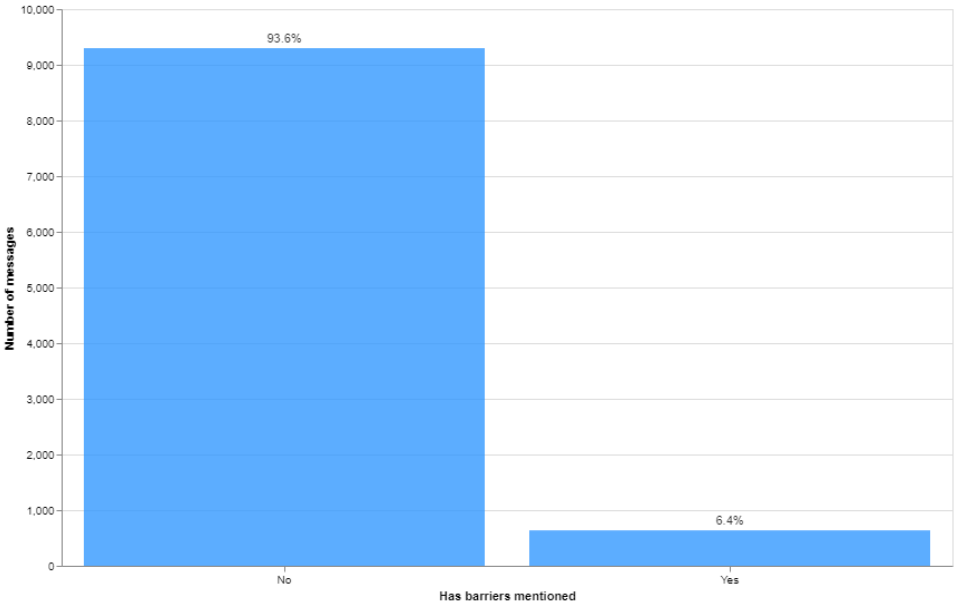
Figure 27: Topic prevalence over time, by month



4.2.2 What barriers are mentioned?

As we did for the general online public, we now identify barriers to inclusive education from the point of view of organisations active in inclusive education landscape in Kazakhstan. Here, we find that, out of close to 10,500 relevant records, about 640 (6.4%) mention barriers to inclusive education (Figure 28). To identify topics mentioned as barriers, we undertake topic modelling on these relevant records. This is discussed next.

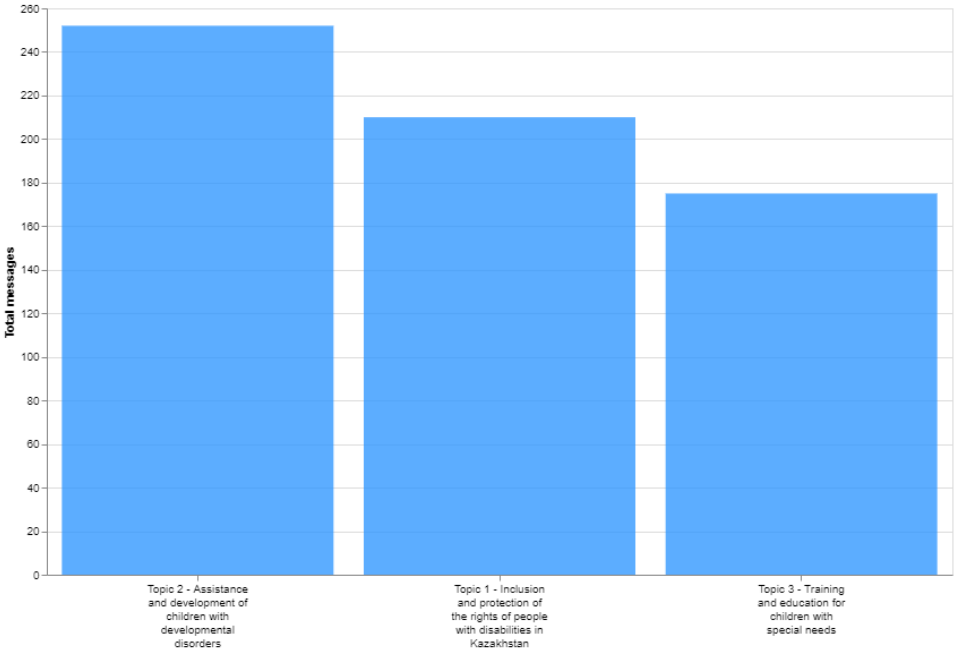
Figure 28: Mentions of rights barriers (Corpus 2)



Topic modelling on the 640 documents yields three main topics of discussion on barriers to inclusive education among organisations working on inclusive education: the need for assistance, such as dedicated centres to address the SEN of children with disabilities;

protection of the rights of children with disabilities; and training of teachers and pedagogical support for children with disabilities (Figure 29). These themes resonate with the topics identified for the overall Corpus 2 regarding disability and education, as discussed in the previous section.

Figure 29: Topics of rights barriers (Corpus 2)



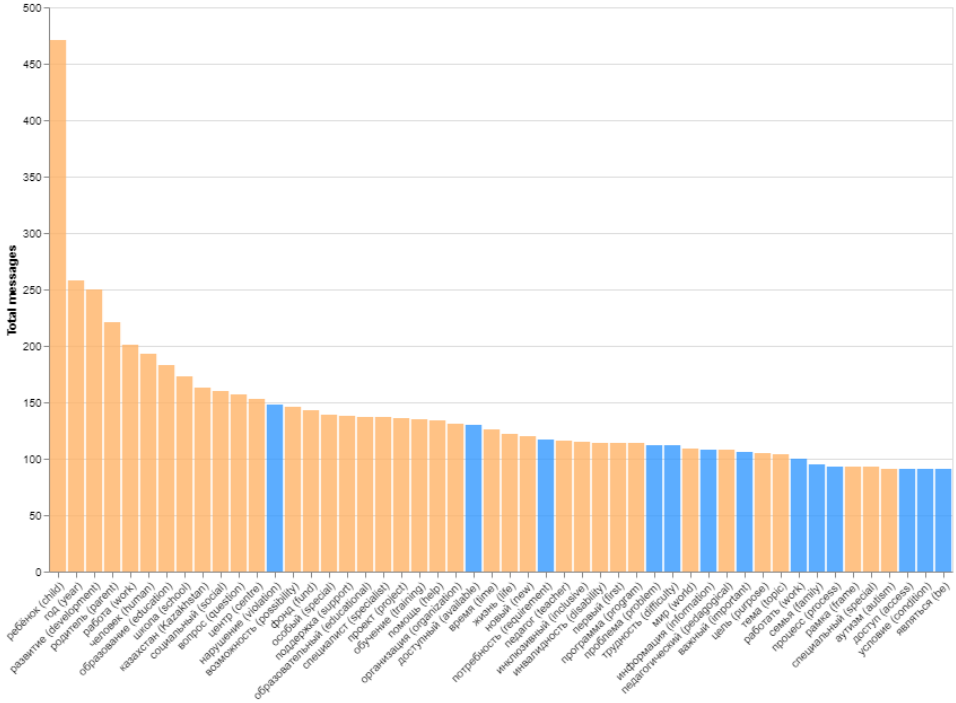
To better comprehend the topics presented in Figure 29, we provide the keywords that feature within these broad topics of discussion (Table 5). As in previous section, we provide English translations for the Russian topics and keywords, which are presented in Annex Table 8.

Table 5: Topic modelling of rights barriers in Corpus 2 (English)

Keywords in each topic	Assistance and development of children with disabilities	Inclusion and protection of the rights of people with disabilities in Kazakhstan	Training and education for children with SEN
1	help	inclusive	special
2	childlike	society	special
3	time	year	course
4	violation	social	pedagogical
5	development	right	program
6	centre	Kazakhstan	teacher
7	year	project	training
8	autism	fund	educational
9	parent	disability	school
10	child	human	education

Next, we look at the most cited words on barriers to inclusive education from these messages and documents to check for terms that have not yet appeared in the analysis on Corpus 2. These terms show that ‘child’, ‘year’, and ‘development’ are the three most common words of discussion (Figure 30).

Figure 30: Most cited words on barriers (Corpus 2)



4.2.3 Are specific programmes being mentioned?

For the third EQ, we did not find mentions of any specific programmes using the automated scraping and analyses offered by Meltwater. To check further for any relevant mentions that may have been excluded from the automated analytics in Corpus 1, we created Corpus 2 using self-programmed scrapers that manually pulled discussions on inclusive education from publicly accessible social media platforms. In our selection of close to 10,000 documents, we found few mentions of any specific programmes. The most common items featuring in this category were the projects ‘Educated Nation’ (Образованная нация) and ‘Comfort School’ (Комфортная школа), which were mentioned in the analysed text corpus.

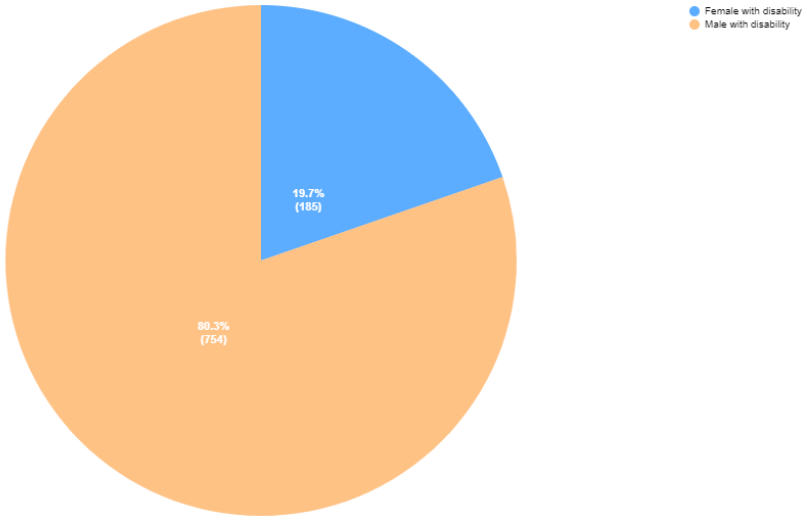
‘Educated Nation’ 2021–25 aims to transform Kazakhstan into a knowledge-based society equipped to meet the challenges and opportunities of the 21st century. The specific programme objectives relating to inclusive education include the aims to facilitate access; to ensure equity in education; to create favourable conditions and an environment for learning; and to improve the quality of education. Meanwhile, the pilot national project in education ‘Comfort School’ aims to eliminate hazardous facilities, to implement three-shift education, and to address the shortage of student seats in secondary education institutions.

This suggests that the discussion on inclusive education overwhelmingly focuses on topics and themes that resonate with the general online public, as well as with organisations working in this space. In comparison, mentions of specific programmes, a focus of EQ 3 of this research, are limited. However, this does not imply that these programmes are not important. It only suggests that, in discussions on inclusive education, issues are discussed more prominently than specific policies or programmes, which are often broad and encompass a variety of aims and objectives.

4.2.4 Gender discussion

As discussed in Section 4.1.3 for Corpus 1, in Corpus 2, also we find that the absolute number of messages, both for education and disability and for only education, is higher for boys than girls (Figure 31). However, the total number of discussions on gender are much smaller here than in Corpus 1. Thus, to avoid any bias emerging from a small sample, we rely on the findings on this theme from Corpus 1.

Figure 31: Gender analysis (Corpus 2)



4.2.5 Summary of key findings

The findings from the analysis of Corpus 2 shed light on key areas of discussion on inclusive education in Kazakhstan among organisations most active in this area in Kazakhstan. Spatial patterns of online conversations show that these are mostly concentrated in large urban cities like Astana and Almaty, even after accounting for higher population densities. This is suggestive of higher engagement on inclusive education in these areas. Social media access may also be a function of economic wellbeing, as a result of which such engagement is lower in other areas of the country.

As seen in results for Corpus 1, the volume of messages rose as the academic session started and tapered off gradually. However, since 2022, there does appear to have been a gradual increase in discussions involving organisations that work in inclusive education. In contrast to the results from Corpus 1, where the online engagement of the general public with inclusive education was higher on news sites, blogs, and other portals, these organisations mostly use platforms like Instagram and Telegram groups for public discussion. Thus, even though general online public and specific organisations working on inclusive education discuss multiple common issues, their preference in voicing these perceptions differs in terms of their platform of choice.

In terms of types of disability, and as seen in Corpus 1, autism was the most discussed. Moreover, there was a significant focus on the capacity building of teachers and specialists who can support the SEN of children. Simultaneously, issues regarding the protection of

rights for children and the importance of early diagnosis and treatment were also widely discussed. The importance of these topics also appears to be deeply entrenched, and there is little variation seen in their relative mentions over time.

The abovementioned themes also featured prominently as major barriers against inclusive education, where the need for dedicated centres to cater to the SEN of children, the need to protect their rights, and improved teacher training and pedagogy were widely discussed. We find few mentions of specific programmes on inclusive education. This is not to say that these programmes are not important, but this indicates that specific issues, rather than all-encompassing programmes, are being discussed more widely. On analysing these results by gender, although we find that the total number of messages on both the themes of education and disability and of education only are higher for boys than for girls, the overall sample for this analysis is small and thus prone to small sample biases. To account for this possibility, we rely on a similar analysis undertaken for Corpus 1, which reveals similar results.

4.3 Our findings in the context of the broader formative evaluation and implications for inclusive education policymaking in Kazakhstan

An important objective of this evaluation was to study the complementarities of this workstream alongside the formative evaluation of the Junction Bulgaria team. Together, these independent studies are expected to generate evidence to inform the inclusive education landscape in Kazakhstan. Despite different evaluation focus and methodological approaches, some findings are mutually reinforcing and are listed below.

First, despite the efforts of the Government of Kazakhstan to broaden the popular understanding of children with SEN, most stakeholders continue to identify this group as children with physical and mental disabilities. Using SML tools, we also find that the most popular discussions on inclusive education are on the rise in reported cases of children with disabilities and measures for its prevention. Analyses of both text corpora reveal that autism and cerebral palsy are widely discussed. This suggests that the medical model, which focuses on characteristics of diseases and disorders, continues to dominate the view of inclusive education. Although discussion on specific conditions is important, a narrow focus on the medical model over the medium- and long-term undermines the social model of disability, which may impede unconditional and equitable access to education, especially for those diagnosed with these conditions.

Second, both studies identify early identification and childhood support for disadvantaged children as an important area of intervention. Our findings regarding the top news stories and the most-mentioned words in Corpus 1 show active discussions on measures taken to prevent disability in newborns and young children, with topics such as equipping perinatal centres to prevent disability featuring commonly in online discussions. The qualitative data used in the formative evaluation also show that the parents of children with SEN are dependent on scarce and expensive options like special kindergartens, private centres, or homeschooling, which can be exclusionary. Our analysis in Corpus 1 finds positive mentions regarding aspects such as autism centres, charity events, and efforts to promote sports participation and improve school accessibility for children with disabilities.

Third, in the context of large geographical distances to reach schools and issues of social stigma around these conditions, we also found that the government recognises the task of addressing early childhood interventions as a major challenge. Our analysis of top news stories among the general online public shows that the Government of Kazakhstan has been taking active steps to introduce modern technologies, creating feedback channels and introducing initiatives aimed at protecting the rights of children with SEN. Moreover, our analysis of 'rights and barriers' reveals that incidents of violence and bullying in kindergartens, schools, universities, and families have been identified by the Government of Kazakhstan as an important area of intervention. We also find evidence of positive news stories on aspects such as accessible school buses, which can tackle inaccessibility arising from large geographical distances.

Fourth, despite these positive steps, the formative evaluation's analysis on heads of education expenditure reveals that a predominant share of funding has recently been prioritised for upgrading facilities like classrooms and buildings. In comparison, lower provisions have been made in favour of teacher training, methodological support, and digitisation. Importantly, it appears that the government is aware of these issues and is moving to address them. We find capacity building of teachers and specialists is an important area of discussion. This topic appears in the top positive news stories in Corpus 1 and is the most-mentioned topic in Corpus 2. Importantly, positive sentiments here may be indicative of steps being taken in the right direction. Analysis of top news stories in Corpus 1 also shows that the Government of Kazakhstan is already taking measures to introduce modern digital technologies to foster inclusive education.

4.4 Caveats, gaps, and limitations

We now turn to discussing the caveats that must be kept in mind while interpreting the above findings. We also state the gaps and limitations that exist in this study and how they can shape a research agenda for the future.

1. This study complements the formative evaluation, which focuses on the role of different state education programmes in promoting inclusive education. Instead of looking at specific government programmes, we provide perspectives of different stakeholders on the inclusive education landscape and conduct a sentiment analysis of these stated opinions by using publicly accessible social media sources.

2. As these perspectives are unknown *a priori* and come from publicly accessible social media data that are different from data collected via traditional survey-based methods, we worked with previously determined EQs but deliberately stayed away from the discussion on effect sizes that emerges from specific hypotheses central to conventional quantitative evaluation studies.

3. For all the benefits that SML tools have to offer, our ability to analyse publicly available information on social media is constrained by socioeconomic factors, as well as by the conscious choice made by individuals and institutions to engage with social media platforms. Thus, the findings emerging from our study, although helpful in understanding multifarious perspectives on inclusive education, are not representative of the country or of any subgroup level. Thus, this stream of research needs to be juxtaposed with studies that lend a voice to those stakeholders of the inclusive education landscape who are unable to air their views on social media platforms.

4. As most social media platforms only keep digital records for a limited period, the analysis is restricted to slightly more than one year. Thus, we are unable to provide insights into how these social media discussions have evolved over the medium- to long-term. As the social media footprint in Kazakhstan grows, this, in our opinion, is an exciting research area that needs support.

5. For reasons discussed earlier, the scope of the study covers children with disabilities in pre-primary, primary, and secondary education institutions in Kazakhstan. However, we believe that, as the first study of its kind in Kazakhstan, this report sets the foundation for broadening this research area in future.

6. Even though the evaluation targets school-age children, we cannot be assured of their participation because children with disabilities often have less access to information and communication technology (ICT) compared to their non-disabled peers. These barriers can be attributed to various factors: for example, devices, software, and platforms may not be designed with accessibility features that cater to the specific needs of children with disabilities; they can be costly; and teachers, parents, and caregivers may not have sufficient knowledge or training on how to incorporate ICT effectively for children with disabilities.

7. Our ability to analyse social media data across geographies depends on whether the data being scaped online are geotagged. As many users online hide or mask their locations, these aspects may be beyond the scope of the study. We nevertheless provide spatial insights from Corpus 2 of the analysis. This comes from organisations working in the inclusive education landscape, which are more likely to share their geolocations than individual stakeholders.

8. Knowing that barriers faced by different groups in accessing inclusive education may differ systematically, we provide some insights into the results through a gender lens. However, our ability to meaningfully disaggregate analyses by different groups of interest using social media data is constrained by the difficulty in identifying the group or individual identity of the account posting social media posts (for example, if a tweet is posted by a man or woman) and by the difficulty in identifying whether a post talks specifically about gender barriers (for example, if a view is posted is general or talks about challenges faced by women or men in accessing inclusive education). These challenges have been pre-empted, which was why no disaggregation as part of the EQs or indicators was committed to at the inception phase.

9. Although the evaluation has taken care to include researchers with Kazakh, Russian, and English skills, data analysis and reporting across three different languages that involve formal and colloquial terminology (and which are impacted by socio-cultural norms) can affect the interpretation of results.

5 Conclusions and recommendations

5.1 Conclusions

The analysis of the social media data of different stakeholders, as undertaken through Corpus 1 and 2, provides valuable learning to advance the knowledge of the inclusive education landscape in Kazakhstan. Lessons learned are listed below and categorised under the relevant EQs.

5.1.1 EQ1: Content of social and online media discussions on inclusive education

1. While there is engagement on vital themes of inclusive education across multiple social media platforms among both the general audience and specific organisations working in this domain, this represents a relatively small share of the overall discussion on education. A lot more could be done to promote such discussions across different stakeholders. Our results show that certain themes may have greater engagement at different points of the year, as in the case of a spike in discussions on autism in April, the month for raising awareness on the issue. This may be leveraged from a policy point of view to facilitate greater cooperation among various stakeholders, promote synthesis of existing evidence, and advance deliberations with the Ministry of Education and other relevant government departments to promote specific policy action.

2. Analysis of general social media discussions show news portals, blogs, and Facebook actively covers topics on inclusive education. Discussions on other social media platforms among this group are, however, relatively limited. In contrast, analysis of social media discussions among organisations active in the inclusive education space shows much higher engagement on Instagram and Telegram groups for public discussions. Thus, when using secondary sources of information such as SML, it is paramount to consider not just different stakeholders but also varied social media platforms to provide a fair representation of different groups of stakeholders.

3. From Corpus 1, which represents discussions among general audience, positive news stories highlighted support for children with autism, prevention of disability in newborns through early diagnosis and treatment, and aspects such as capacity building of teachers and enhanced social protection. Simultaneously, negative news stories focused on the lack of protection of the rights of children with SEN and the need to focus more on the quality of education. Analysis from Corpus 2, in tandem with what we have seen for Corpus 1, identifies common mentions of words like ‘special courses’, which is tied to the capacity building of teachers and specialists who may then be able to support children’s SEN. Thus, in addition to the inclusive physical infrastructure being prioritised by the government, there has been an increased emphasis on the need to support human capital investments.

4. The volume of messages is found to increase as the academic session starts but declines over time, and thus varies with the academic cycle. Moreover, the social media discussions, especially among organisations working on inclusive education, are concentrated in the major cities of Astana and Almaty. This suggests urban areas

have higher engagement with topics on inclusive education and that there is a need to support studies focusing on the inclusive education landscape in those parts of Kazakhstan that have a small digital footprint, especially on social media platforms.

5.1.2 EQ2: Barriers to inclusive education

1. Violence and bullying in what should be seen as safe spaces—the home, kindergartens, schools, and universities—feature as an important barrier to inclusive education. Given the discouragement these instances may have on students, this needs constant monitoring and quick redressal. As our findings reflect, the Government of Kazakhstan, through measures such as establishing minimum safety requirements, is already taking proactive measures to protect children and youth. As implementation of rules, within legal frameworks or otherwise, often presents a continuous challenge these instances need to be closely monitored. Social media platforms can be an important medium for tracking such incidents. These may be further supported by helplines to assist victims and their families seeking time-bound resolution to their complaints.

2. Our findings show concrete efforts to overcome barriers to inclusive education and to promote the mainstreaming of children with SEN. This is being achieved through measures such as improved bus services, sports events, and setting up special centres to help children improve various skills. In this respect, we found evidence of how charity events and community involvement can assist objectives of improving the wellbeing of, and opportunities for, children with SEN. For example, we have found evidence of a positive reception for work undertaken by non-profits such as the Bulat Utemuratov Foundation, which works with autism. This provides a case for potential synergies between the Ministry of Education and non-profits to overcome resource barriers to transcend financial constraints and include specific competencies that some of these organisations may have honed over time.

3. On differentiated barriers to education by gender, we document that, as seen in social media discussions on education, most gendered discussion on inclusive education is dominated by males, with issues relevant to females finding relatively fewer mentions. As viewing each social media message through the gender lens is not feasible, this remains an important theme for future research. We reiterate that these limitations were foreseen at the inception phase where, based on our discussions with UNICEF, it was decided that no disaggregation as part of the EQs could be committed to.

5.1.3 EQ3: Checking for mentions of any state programmes

1. We built a comprehensive list of state programmes to identify mentions of any relevant state programmes that were independently vetted as part of a validation exercise. As Meltwater lacks the flexibility needed to analyse mentions of such programmes, a manual dictionary-based search query was built and executed on Corpus 2 using self-programmed scrapers. It is important to note that, even though we did not find mentions of any specific state programmes, we do document the government's active role in factors furthering improvements in inclusive education in Kazakhstan. Two vital illustrations have been elaborated in Section 4, and further instances include the introduction of modern technologies to create feedback channels, initiatives aimed at protecting the rights of children, and efforts made to address quality issues, improve standards, and establish minimum safety

requirements to reduce incidences of violence and bullying in kindergartens, schools, universities. This does not necessarily mean that specific programmes are not important, but may suggest that discussions on social media mainly pertain to specific issues on inclusive education, the multiple facets of which may not be captured in all-encompassing programmes.

5.2 Recommendations regarding inclusive education in Kazakhstan

The recommendations emerging from the conclusion and lessons learned in the previous subsection are as follows.

1. **Timing and choice of broadcasting platform for key messages:** The analysis of the results for the general online public (Corpus 1) and organisations working in the area of inclusive education (Corpus 2) shows that engagement with the theme of inclusive education rises with start of the academic year and tapers off gradually. Specifically, engagement of the general online public is found to be sensitive to initiatives such as the month of raising awareness about autism in Kazakhstan. Similarly, engagement with the general online public is higher on platforms like news portals and blogs, whereas social media engagement on Instagram and Telegram is higher for organisations working on inclusive education. This information can be leveraged by the Government of Kazakhstan, particularly by the Ministry of Education and the Ministry of Information and Communications, to relay key messages on inclusive education using varied platforms for different stakeholders. Evidence suggests the positive effects of such media messaging on children's education, which is mediated via parental behaviour (Keefer & Khemani, 2014). This may also have larger spillover effects on how society perceives inclusive education, and provides an opportunity to nudge the discussion on inclusive education from a view dominated by the medical model to a view dominated by the social model.

2. **Focus on enhancing accessibility of preschool, primary, and secondary education facilities:** Despite positive perceptions around the Government of Kazakhstan's and the Ministry of Education's increased investment on upgrading school infrastructure and introducing modern technologies, the findings from the Junction Bulgaria report show that there is a requirement for even higher investment in both physical and human infrastructure and resources. Accessibility audits will help prioritise necessary modifications and ensure compliance with accessibility standards so that physical accessibility can be enhanced.

    - For example, wheelchair-accessible classrooms and bathrooms, improved lighting in classrooms and common areas, and enhanced digital accessibility are some areas these audits can focus on. Investment in equipping schools with appropriate assistive technologies can support students with disabilities in accessing educational materials, participating in classroom activities, and effectively engaging with technology. Incorporating Universal Design for Living principles will promote the use of diverse instructional strategies, multiple means of representation, and varied forms of assessment to accommodate different learning styles and abilities. Provision of training and professional development opportunities for teachers and staff on inclusive education practices, disability awareness, and accessibility guidelines will enable educators to better support students with disabilities and create a more inclusive learning environment. Involving parents and caregivers and

establishing open lines of communication to address their concerns will further encourage collaboration on improved accessibility.

Our findings shows that some of these aspects already feature as top positive news stories in Corpus 1 and Corpus 2, which reveals that the government is aware and moving to address these challenges. We propose scaling up these action points to fast-track these gains. Certain voluntary initiatives conducting accessibility audits for public facilities have already emerged—notably Accessible Kazakhstan—and could provide guidance for the Government of Kazakhstan to institutionalise this model in the context of educational facilities.

3. **Investment in human capital:** Based on repeated mentions of teacher training and skill upgradation, we recommend providing training and professional development opportunities for teachers and staff members on inclusive education practices, disability awareness, and accessibility guidelines to enable educators to better support students with disabilities and create a more inclusive learning environment. Training could also contribute towards a more individual approach being taken to students with SEN. Involving parents and caregivers, and establishing open lines of communication to address their concerns, will further encourage collaboration on improved accessibility. These training programmes should include modules on the identification and prevention of bullying and violence against children with SEN, and should focus on building inclusive language to minimise the use of potentially stigmatising vocabulary.

4. **Coverage of social media dark areas:** As acknowledged previously, we would like to state again that the social media data on which this analysis is based come from larger cities like Astana and Almaty. However, more research is needed to understand perceptions of, and barriers to, inclusive education in other parts of the country. This can then guide policy action that would be best suited to advancing the inclusive education goals and aspirations of Kazakhstan's population.

5. **Gender-sensitive assessment of educational needs:** Our results show that girls are relatively underrepresented in online discussions on both education and inclusive education in Kazakhstan, suggesting a lack of attention paid to the differential needs of girls in the educational system. This is in tandem with Junction Bulgaria's formative evaluation findings, which show that state education programmes do not mention measures taken to reduce gender disparities in educational outcomes, not least because of the scarcity of available data. This calls for a collection of indicators disaggregated by gender within the education sector in order to identify gender-specific factors limiting access to inclusive education, including the potential impact of early marriage and teenage pregnancies. Such an analysis could inform more gender-sensitive policies within the inclusive education sector.

6. **Early childhood interventions:** Early prevention and diagnosis of disabilities, as well as the provision of services for early childhood development for children with SEN, are crucial to ensure that their educational needs are addressed from the outset. Positive news stories online indicate some progress has been made in measures taken to prevent disability in newborns and young children, including through better equipping perinatal centres. These initiatives should be expanded in coverage and coupled with correctional support to prepare young children with SEN to study in public schools. Positive news stories also mention the role of social protection programmes, which could be used to cover disability-related costs, as well as any expenditure on correctional support.

7. **Adaptation of standardised testing:** The EHT hashtag appeared commonly in our analysis, showing active discussions around the need to adapt standardised tests to the needs of children with SEN. Our findings from top news stories in Corpus 1 indicate that the Government of Kazakhstan has provided an additional 40 minutes for children with SEN during such tests, showing that progress on this issue is already underway. Further adjustments could be made to adapt testing materials, for instance for children who cannot see the text. Once those adjustments have been made, the government should provide information for parents of children with SEN to inform them of their rights during the test and motivate them to pursue the EHT, which is a key pathway into college and university education.

5.3 Recommendations for the future use of SML in evaluations

The rising volume of social media discussions presents a unique opportunity to leverage publicly available Big Data sources of information in a variety of evaluation contexts. At the same time, this innovative approach differs from traditional survey-based evaluation methods and carries its own set of considerations. Based on our methodological approach, the following recommendations for future use of SML in evaluations emerged.

1. **Construct EQs carefully and be conscious of what questions SML can and cannot answer:** Traditional impact evaluations using quantitative data are often concerned with causal questions, such as estimating the causal impact of an intervention on a set of outcomes. SML is better suited to address questions that are more descriptive in nature and are intended to capture the perspectives of stakeholders on a specific topic of online discussion. Examples of questions SML can answer include: 1) Landscaping: What is the online debate in-country about a specific topic? 2) Actor analysis: What is the online influence of the stakeholder group of interest? 3) Longitudinal study: How has the debate on a topic changed over time?

2. **Validation of data pulled from social media is key to ensuring relevance:** The process of building search queries should be iterative and based on expert discussions. The set of keywords needs to take into account any linguistic or cultural sensitivities, while being simple enough to ensure a sizeable corpus. Once a search query is built, results should be analysed by experts and changed iteratively until a satisfactory level of relevance is reached. Note that this process can often be very sensitive, as one word can skew the analysis significantly. Thus, we recommend running multiple versions of the search query on a particular topic.

3. **Exclusion and inclusion biases must be accounted for:** Depending on the evaluation context, inequalities in access to social media can result in certain groups being excluded from online discussions, often limiting the potential of studies using SML tools to be representative at the national, or even regional, level. The social media coverage in-country, as well as factors potentially limiting access, should be considered to identify an appropriate stakeholder group whose online discussions will be analysed. Bias can also arise from the fact that people tend to be more extreme on social media than in person and are more likely to engage with controversial material. This further highlights the need for expert discussions to validate the relevance of online posts.

4. **External validity should be considered:** Given that online discussions are not necessarily representative of the entire population and that in some cases the number of posts is small, the evaluation should clearly state whether the results can be generalisable. While an online debate on a particular topic is unlikely to be a true

representation of an actual debate in-country, it is an interesting and complementary analysis that offers additional insights that are not captured by traditional surveys.

5. **Meltwater vs self-programmed scrapers:** There are advantages and disadvantages to the use of online platforms relative to self-programmed scrapers, depending on the stakeholder group and the EQ at hand. While SML tools such as Meltwater cover a wider range of social media sources and allow capturing discussions of the general public, their analytical tools can be limited, for instance the inability to download raw social media data and carry out independent analysis. On the other hand, self-programmed scrapers allow us to access discussions of a specific stakeholder group, such as a Telegram chat for organisations working in the inclusive education space in Kazakhstan, and enable us to conduct our own NLP analysis, although this needs to be written for every social media source separately.

6 References

Allan, J., and Omarova, T. (2021) Disability and inclusion in Kazakhstan, *Disability & Society* 37(7). doi:https://doi.org/10.1080/09687599.2020.1867073

Grimmer, J., and Stewart, B.M. (2013) Text as data: The promise and pitfalls of automatic content analysis methods for political texts, *Political Analysis* 21(3), pp. 267–297.

Junction Bulgaria (2023) 'Formative Evaluation of the State Programmes for Development of Education with the Focus on Inclusivity', Sofia.

Keefer, P., and Khemani, S. (2014) Mass media and public education: The effects of access to community radio in Benin, *Journal of Development Economics* 109, pp. 57–72.

Marteau, J.-F. (2020) 'Post-COVID Education in Kazakhstan: Heavy Losses and Deepening Inequality', *World Bank Blogs. Eurasian Perspectives* (online), accessed at https://blogs.worldbank.org/europeandcentralasia/post-covid-education-kazakhstan-heavy-losses-and-deepening-inequality.

Ministry of Education (2023) 'National Report on the State and Development of the Education System KZ', Government of Kazakhstan, Astana.

Sustainable Development Report (2023) 'Sustainable Development Report (2023) Kazakhstan, Indicators' (online), accessed at https://dashboards.sdgindex.org/profiles/kazakhstan/indicators.

UNESCO, UNICEF, and FCDO (2021) 'Education Sector Analysis Methodological Guidelines, Volume III'.

UNICEF (2019) 'Situational Analysis of Children in Kazakhstan', UNICEF (online), accessed at www.unicef.org/kazakhstan/media/4611/file/Situation%20Analysis-2019.pdf.

Annex A Terms of reference

PART I		
Purpose of Assignment	**Inclusive education evaluation by using Big data source of information (descriptive**	
Location of Assignment	Remote	
Duration of contract	7 months	
Start period (estimated)	From: December 2022	To: 31 July 2023
Reporting to:	Raushan Ibrasheva, CRM Specialist	
Budget Expiry Date	**31 December 2022**	

A.1 Objectives, Purpose & Expected Results

The purpose of the formative evaluation is to examine the public opinion (parents/ caregivers/ teachers) on to what extent the state education programmes in Kazakhstan implemented within the period of 2011-2021 have been contributing to inclusion of children of different levels of ability in the education system in the country as well as to examine the enabling conditions and bottlenecks which will require further improvement. Internet public opinion can be disseminated by e-mail, forum, blog and other means of social media interactions. Public opinion information appears in the form of network comments, but network comments include spam comments, subjective comments and objective comments. The classification of these comments needs to extract the user's feelings. Subjective comments involve a large number of user ideas, rather than subjective comments containing incorrect opinions and lack of emotional statements. It is meaningful to get subjective comments from a large number of online comments. This particular assignment will serve as a good addition to the formative evaluation and provide the Government of Kazakhstan and UNICEF with sound evidence and conclusions to inform planning the future strategies and work in this area.

Within this framework, the evaluation will have the following specific objectives:

1. Through the analysis of the perspective of individuals, in particular parents, teachers and youth and overall public opinion

1.1. Assess some of the immediate results and good practices of what is working, for whom, and what is not working, from the perspective of parents and general public opinion in the inclusive education.

1.2. Identify existing challenges, barriers and gaps in the education system and provide strategic recommendations aimed at creating an inclusive and resilient learning environment for every child.

1.3. develop recommendations for future strategies in sustaining results and good practices and addressing existing challenges and barriers.

A.2 Deliverables

A.2.1 Inception Report

The team will present the methods to be used for this work, the scope, project plan with timelines, the triangulation/validation methods. The research team responsible for this assignment should be part of existing reference group to actively coordinate and engage during presentations at a workshop. The workshop will be delivered with participation of research team responsible for formative evaluation, reference group, programme staff and external stakeholders. This will include validation or rebuttal of the recommendations by the stakeholders.

A.2.2 Digital Ecosystem Analysis

1. Technical Report should include:

- Analysis of the perspective of individuals, in particular parents, teachers and youth and overall public opinion, which is consist of some of the immediate results and good practices of what is working, for whom, and what is not working, identifying existing challenges, barriers and gaps in the education system and provide strategic recommendations aimed at creating an inclusive and resilient learning environment for every child, with recommendations for future strategies in sustaining results and good practices and addressing existing challenges and barriers.

2. A Digital Ecosystem Analysis Report:

- Size and demographic breakdown of conversations around inclusive education.
- Discourse across platforms and how each platform is used.
- Media communication approaches about inclusive education.
- Technologically related deliverables: data used for the analysis and scripts for the categorization
- Methodological note

A.2.3 Deliverable #3: Final package with methodology and tools and capacity building

1. Full Methodological note:

- The note can be adopted by other countries outlining the key requirements to enable adoption (data availability, workshop processes, etc)

2. Data architecture:

- Including algorithms used to captures, analyse, and utilise information
- An analytical tool that provides and enables a near real time population level view of current discourse on inclusive education that translates outputs in a meaningful way to the end-users. The tool should allow monitoring data every quarter. The tool

should include social and search data that breaks down demographic skews and overall discourse and trends around inclusive education. It should include the digital ecosystem analysis for the Kazakhstan and should aim at expanding to other countries. It should serve as baseline of how public discourse of inclusive education perception changed over time to allow monitoring of the effectiveness and reach future inclusive education programmes/campaigns.

3. Manual:

- For the tool on the steps ahead to make it sustainable

4. Capacity building online workshop:

- To ensure that UNICEF Kazakhstan implementing partners (Ministry of Education) and relevant staff are equipped to leverage digital data for programs and policies based on this particular assignment. The workshop should provide training and manuals and training materials.

Annex B : Additional Figures and Tables

Figure 32: Building a search query in Meltwater (KZ)

Figure 33: Building a search query in Meltwater (RU)

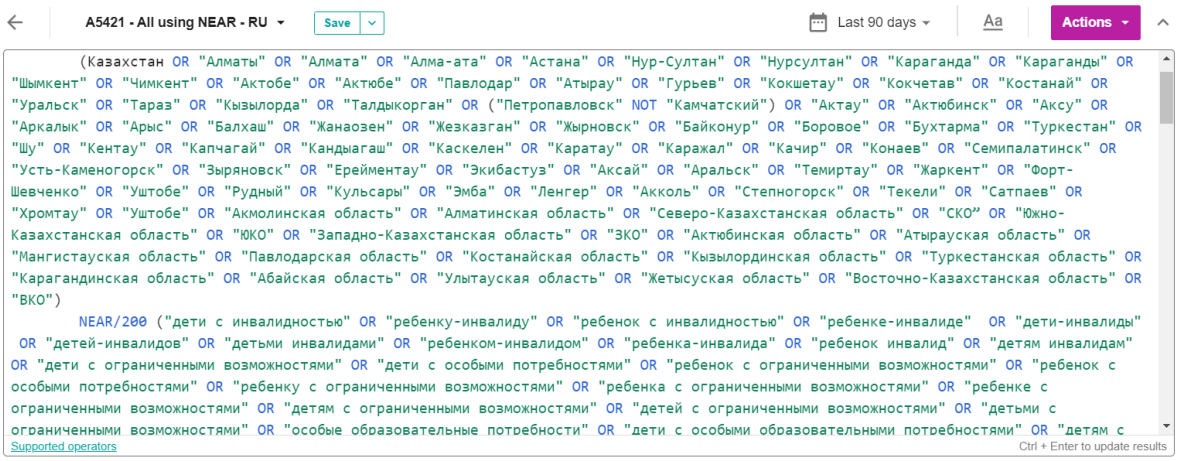


Table 6: Corpus 2, List of Organisations

WEB	TELEGRAM
https://bilimland.kz/ru/news-articles/news	Dara Charity Foundation
https://www.cpm.kz/ru/news	Mama Pro Kazakhstan кz
https://www.kzhol.kz/poslednie_novosti	Про Инклюзию | Инклюзия Жайлы (both group and chat)

https://orleu-edu.kz/ru/orleunews/	ПРАВА МЕНТАЛЬЩИКОВ
https://special-edu.kz/news/6	
https://utemuratovfund.org/news	
https://bilimland.kz/ru/news-articles/news	

INSTAGRAM CHANNELS		
https://www.instagram.com/ardi_kz/	https://www.instagram.com/dara.inclusion/	https://www.instagram.com/pmpk_karaganda/
https://www.instagram.com/ario_kz_/	https://www.instagram.com/doskaz.kz/	https://www.instagram.com/prava_ocobennogo_rebenka/
https://www.instagram.com/asylmiras_astana/	https://www.instagram.com/inclusionteam/	https://www.instagram.com/ravmir_astana/
https://www.instagram.com/bolashakcharity/	https://www.instagram.com/juldyzai.fond/	https://www.instagram.com/specialedu.kz/
https://www.instagram.com/bulat_utemuratov_foundation/	https://www.instagram.com/kzhol_charityfund/	https://www.instagram.com/spectrum.astana/
https://www.instagram.com/centerbalama/	https://www.instagram.com/nuriya.baltabaikyzy/	https://www.instagram.com/veneraclub.atyrau/
https://www.instagram.com/centerkenes/	https://www.instagram.com/okoo.kz/	https://www.instagram.com/zeiin_atyrau/
https://www.instagram.com/centre.davinci/	https://www.instagram.com/orda_autism/	https://www.instagram.com/zhasurpaq_nur/
https://www.instagram.com/cpm_official_page/	https://www.instagram.com/orleu.edu.kz/	https://www.instagram.com/pmpk_karaganda/
https://www.instagram.com/dara.charity/	https://www.instagram.com/pmpk1_astana/	https://www.instagram.com/prava_ocobennogo_rebenka/

Table 7: Topic modelling (Russian)

Code	Topic description	Keywords									
1	Развитие потенциала школьных учителей	организация	год	развитие	работа	область	педагогический	образование	учитель	школа	педагог
2	Содержание школьного образования	получить	занятие	школа	развитие	тема	повышение	класс	обучение	программа	курс
3	Важность образования в жизни человека	хороший	занятие	важный	жизнь	время	нарушение	развитие	человек	родитель	ребёнок
4	Педагогическая поддержка и работа специалистов	провести	поддержка	результат	пройти	педагогический	работать	год	работа	специалист	центр

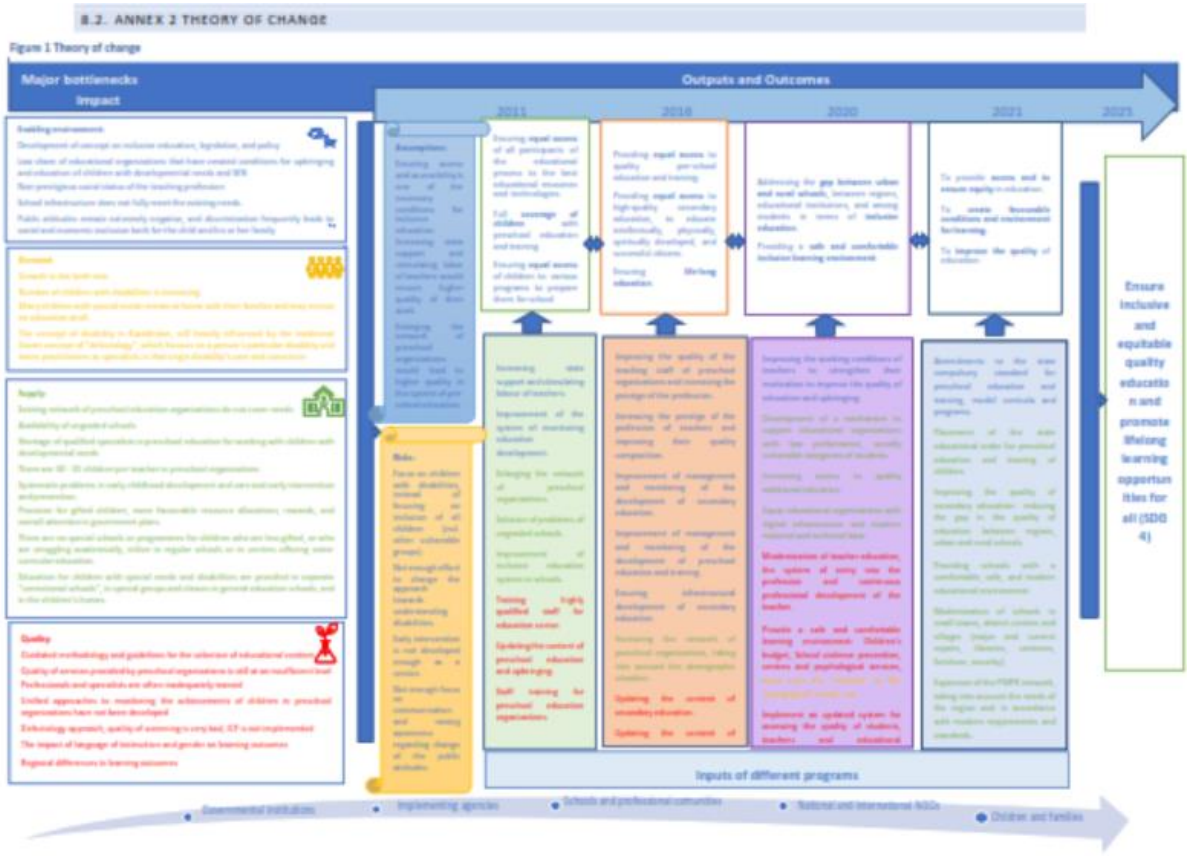
Code	Topic description	Keywords									
5	Особенности образования детей с потребностями	организация	педагогический	школа	потребность	образовательный	особый	инклюзивный	специальный	ребёнок	образование
6	Диагностика и лечение детей с особыми потребностями	получить	пройти	новый	детский	вопрос	ребёнок	диагноз	инвалидность	аутизм	год
7	Возможности социальной поддержки и проекты в образовании	рамка	возможность	мир	город	год	поддержка	казахстан	социальный	фонд	проект

Table 8: Topic modelling of Rights Barriers in corpus 2 (Russian)

Code	Topic description	Keywords									
1	Инклюзия и защита	инклюзивный	общество	год	Социальный	право	казахстан	Проект	фонд	инвалидность	человек

Code	Topic description	Keywords									
	прав людей с инвалидностью в Казахстане										
2	Помощь и развитие детей с нарушениями развития	помощь	детский	время	нарушение	развитие	центр	год	аутизм	родитель	ребёнок
3	Обучение и образование для детей с особыми потребностями	особый	специальный	курс	Педагогический	программа	педагог	Обучение	образовательный	школа	образование

Figure 34: Theory of Change (Junction Bulgaria)



Reference: Junction Bulgaria (2022). Formative Evaluation of State Programmes for Development of Education with Focus on Inclusivity. Inception Report, pp. 48.