

Б.С. Абдрасилов , Ш.Б. Алтыбаева , Г.С. Ерсултанова* ,
И.Ж. Жуматова , К.Т. Турадбаева 

Национальный центр тестирования, Астана, Казахстан

*e-mail: ersultanovagulnur@gmail.com

ИНТЕГРАЦИЯ ЦИФРОВЫХ ПЛАТФОРМ В РАЗРАБОТКУ ТЕСТОВ ДЛЯ НАЦИОНАЛЬНЫХ ЭКЗАМЕНОВ: ВЫЯВЛЕНИЕ ТЕНДЕНЦИЙ И РЕКОМЕНДАЦИИ

В исследовании проанализирован потенциал цифровой трансформации процессов разработки тестовых заданий для национальных экзаменационных систем Казахстана (ЕНТ, КТ, КАЗТЕСТ) и выявлены ключевые ограничения текущих практик. Несмотря на внедрение компьютерного тестирования и отдельных цифровых решений (прокторинг, электронные сертификаты), критически важный этап создания и экспертизы заданий остаётся преимущественно ручным, что снижает эффективность, масштабируемость, валидность и безопасность тестов. На основе комплексного анализа нормативной базы и эмпирических данных обоснована необходимость внедрения специализированной цифровой платформы, автоматизирующей весь жизненный цикл тестового задания – от генерации с использованием когнитивных моделей и шаблонов (AIG) до экспертизы, корректировки, вариативности и аналитики.

Отмечается соответствие предлагаемого решения международным стандартам (AERA, APA, NCME) и требованиям Болонского процесса. Несмотря на высокую готовность вузов к внедрению (80,5%), выявлены основные барьеры: недостаточная техническая инфраструктура, финансовые ограничения и организационное сопротивление. Подчёркивается стратегическая значимость платформы для повышения качества, прозрачности и доверия к системе оценивания в условиях цифровой трансформации образования. В опросе приняли участие 206 респондентов, давших согласие принять участие в исследовании. Эмпирическое исследование в виде опроса охватывало 22 вуза Казахстана.

Статья будет полезна разработчикам образовательной политики, специалистам в сфере тестирования, ИТ-разработчикам, исследователям, а также представителям вузов и международных организаций, интересующимся модернизацией систем оценивания.

Ключевые слова: национальные экзаменационные системы, цифровая трансформация образования, автоматизация разработки тестовых заданий, валидность и надежность тестирования, цифровая платформа оценивания, автоматическая генерация заданий.

B.Abdrasilov, Sh. Altybayeva , G. Yersultanova*,
I. Zhumatova, K.Turalbayeva

National Testing Center, Astana, Kazakhstan

*e-mail: ersultanovagulnur@gmail.com

Integration of Digital Platforms into the Development of Tests for National Exams: Identifying Trends and Recommendations

The study analyzes the potential of digital transformation of the processes of developing test assignments for the national examination systems of Kazakhstan (UNT, CT, KAZTEST) and it identifies the key limitations of current practices. Despite the introduction of computer-based testing and individual digital solutions (proctoring, electronic certificates), the critical stage of creating and examining assignments remains predominantly manual, which reduces the effectiveness, scalability, validity and security of tests. Based on a comprehensive analysis of the regulatory framework and empirical data, the need for the introduction of a specialized digital platform automating the entire life cycle of a test assignment – from generation using cognitive models and templates (AIG) to expertise, adjustments, variability and analytics.

The compliance of the proposed solution with international standards (AERA, APA, NCME) and the requirements of the Bologna process is noted. Despite the high readiness of universities for implementation (80.5%), the main barriers were identified: insufficient technical infrastructure, financial constraints and organizational resistance. The strategic importance of the platform for improving the quality, transparency and trust in the assessment system in the context of the digital transformation of education is

emphasized. The survey involved 206 respondents who agreed to participate in the study. The empirical study in the form of a survey covered 22 universities in Kazakhstan. The article will be useful for educational policy makers, testing specialists, IT developers, researchers, as well as representatives of universities and international organizations interested in upgrading assessment systems.

Keywords: national examination systems, digital transformation of education, automation of test item development, validity and reliability of testing, digital assessment platform, automatic item generation (AIG)

Б.С. Абрасилов, Ш.Б. Алтыбаева, Г.С. Ерсултанова*,
И.Ж. Жуматова, К.Т. Туралибаева

Ұлттық тестілеу орталығы, Астана, Қазақстан

*e-mail: ersultanovagulnur@gmail.com

Ұлттық емтихандарға тест өзірлеуде цифрлық платформаларды интеграциялау: үрдістерді анықтау және ұсыныстар

Зерттеуде Қазақстандағы ұлттық емтихандық жүйелердің (ҰБТ, КТ, ҚАЗТЕСТ) тест тапсырмаларын өзірлеу үдерістерін цифрлық трансформациялау өлеуеті талданып, қолданыстағы тәжірибелердің негізгі шектеулері айқындалды. Компьютерлік тестілеу мен жекелеген цифрлық шешімдердің (прокторинг, электронды сертификаттар) енгізілуіне қарамастан, тапсырмаларды құрастыру мен сараптау сияқты шешуші кезеңдер басым түрде қолмен орындалады, бұл тестілердің тиімділігіне, ауқымдылығына, валидтілігіне және қауіпсіздігіне теріс әсерін тигізуде. Нормативтік базага және эмпирикалық деректерге кешенді талдау жасау негізінде тест тапсырмасының өмірлік циклін толық автоматтандырылған цифрлық платформаны енгізуін қажеттілігі дәлелденді. Мұндай платформа когнитивтік модельдер мен үлгілерді пайдалана отырып автоматтандырылған тапсырма генерациясынан (AIG) бастап сараптау, түзету, вариативтілік пен аналитикаға дейінгі барлық кезеңдерді қамтиды.

Ұсынылып отырган шешімнің халықаралық стандарттарға (AERA, APA, NCME) және Болон процесінің талаптарына сәйкестігі атап өтіледі. Жоғары оқу орындарының енгізуге дайындығы (80,5%) жоғары болғанымен, негізгі кедергілер ретінде техникалық инфрақұрылымның жеткіліксіздігі, қаржылық шектеулер және ұйымдастырушылық қарсылықтар анықталды. Цифрлық трансформация жағдайында бағалау жүйесінің сапасын, ашықтығын және сенімділігін арттыру үшін платформаның стратегиялық маңызы ерекше атап көрсетіледі. Саяуламаға зерттеуге қатысуға келісім берген 206 респондент қатысты. Саяулама түріндегі эмпирикалық зерттеу Қазақстанның 22 жоғары оқу орынның қамтыды. Мақала білім беру саясатын өзірлеушілерге, тестілеу саласының мамандарына, IT-өзірлеушілерге, зерттеушілерге, сондай-ақ бағалау жүйесін жаңғыру тұға мүдделі жоғары оқу орындары мен халықаралық ұйымдардың өкілдеріне пайдалы болады.

Түйін сөздер: ұлттық емтихан жүйелері, білім беруді цифрландыру трансформациясы, тест тапсырмаларын өзірлеуді автоматтандыру, тестілеудің валидтілігі мен сенімділігі, бағалаудың цифрлық платформасы, тапсырмаларды автоматтый генерациялау.

Введение

Цифровая трансформация стала ключевым фактором модернизации всех сфер общественной жизни, включая образование. Глобальные вызовы, такие как пандемия COVID-19, ускорили переход к цифровым технологиям в обучении, преподавании и оценке знаний (UNESCO, 2020). В Казахстане этот процесс затронул национальные системы тестирования, включая переход на компьютерный формат проведения Единого национального тестирования (далее – ЕНТ), Комплексного тестирования (далее – КТ) и ҚАЗТЕСТ (zakon.kz, 2021). В рамках Концепции развития высшего образования и науки на 2023–2029 годы одним из приоритетных направлений определено совершенствование систем

оценивания через внедрение цифровых технологий (Концепция, 2023). Модернизация процедур оценивания имеет стратегическое значение, так как качество национальных экзаменов зависит от эффективности разработки, управления и цифровизации материалов.

Разработка тестовых заданий для национальных экзаменов требует высокой квалификации экспертов и современных технологических решений, обеспечивающих эффективность, валидность и масштабируемость процессов. Традиционные подходы, основанные на ручной работе экспертов, характеризуются высокой трудоемкостью, фрагментарностью и ограниченными возможностями для стандартизации и коллективной работы. Несмотря на внедрение компьютерного тестирования и онлайн-прокторинга,

процесс создания заданий по-прежнему зависит от текстовых редакторов, что снижает операционную эффективность и затрудняет обеспечение валидности и надежности тестов.

Обзор литературы

Ключевые исследования формируют теоретическую основу настоящего анализа. Так, Гирл и соавт. (2023) подчеркивают роль автоматической генерации заданий (AIG) на основе шаблонов и машинного обучения, что позволяет создавать крупные банки заданий с минимальной субъективностью, адаптированные к двуязычным требованиям казахстанских тестов. Гирл и соавт. (2023) развивают модель AIG, интегрирующую когнитивные модели, шаблоны заданий и алгоритмы генерации, которые варьируют элементы заданий (например, числовые значения или контекст задачи) при сохранении психометрических характеристик, таких как сложность и дискриминационная способность.

Современные подходы к психометрике, включая применение цифровых технологий в тестировании, подчёркиваются в работе Кардановой и Ивановой (2023), что особенно актуально для автоматизации разработки заданий в казахстанских экзаменационных системах (ЕНТ, КТ, КАЗТЕСТ). Брусиловский и Миллер (1999) акцентируют внимание на жизненном цикле тестовых заданий, подчеркивая необходимость автоматизации этапов подготовки, представления и аналитической обратной связи с использованием веб-инструментов, что соответствует задачам модернизации казахстанских систем тестирования. С учётом национального контекста вопросы разработки тестовых заданий нашли дальнейшее отражение в работах казахстанских исследователей. В монографиях Сагиндикова И.У. (2021, 2023) всесторонне рассматриваются методологические и содержательные основы проектирования заданий для ЕНТ, что формирует прочную теоретическую платформу для внедрения инновационных решений, включая цифровизацию процессов и использование автоматизированных технологий создания тестов. Эти подходы формируют концептуальную рамку исследования, направленного на разработку и внедрение цифровой платформы для автоматизации создания тестовых заданий, обеспечивающей соответствие международным стандартам (AERA, APA, NCME, 2014) и требованиям Болонского процесса.

В этом контексте особое значение приобретает международная практика внедрения цифровых решений в процессы тесторазработки. Зарубежный опыт подтверждает высокую результативность использования специализированных платформ, таких как GradeMaker Pro, TAO и собственные разработки Cambridge Assessment, обладающих значительным трансформационным потенциалом в сфере управления банками тестовых заданий (taotesting.com, grademaker.com). В частности, в Великобритании наглядным примером служит деятельность крупнейшей экзаменационной организации AQA (Assessment and Qualifications Alliance), которая в 2023 году интегрировала платформу GradeMaker Pro в процедуры создания экзаменационных материалов. В результате более двух третей заданий для экзаменов GCSE и A-level формируются с её использованием, что демонстрирует высокую эффективность цифровых решений в условиях функционирующей системы массовой внешней оценки (AQA, 2023).

Многие исследования сосредоточены на вопросах валидности и надежности тестов, однако разработка самих заданий часто остается прерогативой экспертов, работающих без интеллектуальных средств автоматизации. Традиционный подход, основанный на компьютеризации ручных процессов, предоставляет возможности для управления банком заданий и автоматической проверки результатов, но не решает проблему трудоемкости их создания. Для защиты от недобросовестного прохождения тестов необходима крупная база заданий и динамическая генерация вариантов, что делает автоматизацию критически важной. В казахстанском контексте переход к компьютерному формату национальных экзаменов был впервые реализован в 2020 году при проведении КТ (Алтыбаева & Скабаева, 2025). Этот процесс сопровождался внедрением методологических принципов создания заданий и организационно-управленческих аспектов интеграции цифровых технологий.

Целью настоящего исследования является анализ и оценка возможности внедрения цифровой платформы для автоматизации формирования тестовых заданий в национальных экзаменационных системах Казахстана. Исследование направлено на формулирование практических рекомендаций по созданию такой платформы, что позволит повысить эффективность, прозрачность и качество разработки заданий, а также обеспечить соответствие требованиям Болон-

ского процесса и международным стандартам образовательного оценивания. В рамках исследования представлены результаты опроса среди представителей казахстанских вузов, направленного на изучение их опыта в области автоматизации процессов создания тестовых заданий.

Материалы и методы исследования

Исследование основано на дизайне встроенного кейс-стади, направленного на изучение автоматизации разработки тестовых заданий в Национальном центре тестирования (далее – Центр) с акцентом на внедрение цифровой платформы для национальных экзаменационных систем. Исследование использовало конвергентный параллельный смешанный дизайн ((Кресвел и Плано Кларк, 2017). Это позволило одновременно собирать и интегрировать качественные и количественные данные, обеспечивая всесторонний анализ процессов цифровой трансформации.

Сбор данных осуществлялся с использованием двух основных методов: анализа документов и опроса. Анализ документов охватывал изучение регламентов Центра, связанных с разработкой тестовых заданий, отчётов об эффективности автоматизированной системы управления базой заданий, а также обзоры программных средств, применяемых при сопровождении апелляционного процесса.

Опрос проводился среди представителей, ответственных за организацию и проведение экзаменов в казахстанских вузах, с целью оценки текущего уровня автоматизации, выявления проблем и определения потребностей в цифровых инструментах. Генеральная совокупность включала 302 специалиста. Учитывая специфику экспертной группы и необходимость обеспечения достаточного охвата, была сформирована целевая выборка, в которую вошли 206 респондентов, давших полное согласие на участие.

При формировании анкеты особое внимание уделялось её структурированности и категоризации вопросов, как это рекомендуется в литературе по опросным методам (Dillman, 2014). Анкета была разработана таким образом, чтобы минимизировать время её заполнения, что, согласно исследованиям, положительно влияет на уровень отклика. Анкета включала 22 вопроса, сгруппированных в 6 тематических блоков.

Надёжность и валидность анкеты обеспечивались проведением пилотного исследования (pre-testing) с участием 12 экспертов, не вошед-

ших в основную выборку. На основе их обратной связи были внесены корректировки в формулировки двух вопросов для повышения ясности и однозначности.

Сбор данных проводился в период с 06 мая по 15 июня 2025 года.

Для статистической обработки количественных данных опроса использовались методы описательной статистики (частотный анализ, средние значения, стандартное отклонение). Для проверки нормальности распределения данных по шкальным переменным был применён критерий Шапиро-Уилка. Результаты показали значимые отклонения от нормального распределения для анализируемых переменных ($p < 0.05$). Учитывая порядковый характер данных (шкала Лайкерта) и отсутствие нормальности распределения, для корреляционного анализа был выбран коэффициент корреляции Спирмена.

Интеграция качественных и количественных данных осуществлялась на этапе интерпретации. Темы, выявленные в результате опроса, были триангулированы с данными анализа документов и статистическими результатами, что обеспечило целостное понимание процесса внедрения цифровой платформы. Такой смешанный методологический подход позволил учесть операционные, организационные и контекстуальные особенности автоматизации разработки тестовых заданий в национальных экзаменационных системах Казахстана.

Результаты и обсуждение

В исследовании была использована полустандартизированная анкета, разработанная с целью выявления актуальных практик, проблемных зон и институциональных потребностей, связанных с процессами разработки экзаменационных материалов в казахстанских высших учебных заведениях. Анкета составлена с учётом рекомендаций по цифровизации образовательной оценки и адаптирована к локальному контексту.

Структура анкеты включает 6 тематических разделов и содержит как закрытые, так и полуоткрытые вопросы:

- Общие сведения: тип и профиль вуза, а также профессиональная роль респондента (вопросы 1–3).

- Процесс разработки экзаменационных материалов – степень автоматизации, использова-

ние шаблонов, этапы согласования и типичные сроки подготовки (вопросы 4–9).

- Использование цифровых технологий – наличие специализированных платформ, уровень цифровизации (вопросы 10–12).

- Интеграция и автоматизация – меры безопасности, применение алгоритмов генерации заданий, учет таксономии Блума, потребность в автоматизации (вопросы 13–17).

- Опыт и потребности в цифровой платформе – оценка важности функционала будущей системы, готовность вуза к её внедрению и существующие барьеры (вопросы 18–21).

- Дополнительные комментарии – открытое поле для респондентов (вопрос 22).

В рамках исследования был проведён комплексный анализ корпуса документов, составляющих нормативную и операционную основу процесса формирования базы тестовых заданий. Эмпирической базой послужили нормативно-регламентирующие документы, в частности «Правила формирования базы тестовых заданий» (далее – Правила), техническая документация, включая спецификации к государственным заданиям услуг по разработке и формированию базы тестовых заданий, внутренние отчёты и материалы служебного документооборота, отражающие фактические операционные процессы и возникающие проблемы.

Многоаспектный анализ позволил реконструировать формализованную модель жизненного цикла тестового задания (*de jure*) и выявить её фактическую реализацию (*de facto*). Это также помогло идентифицировать ключевые разрывы и технологические барьеры, снижающие общую эффективность системы.

Анализ Правил свидетельствует о наличии формализованного и многоэтапного процесса разработки тестовых заданий, включающего составление спецификаций, двух экспертизу, корректировку, апробацию и внедрение в автоматизированную систему управления базы тестовых заданий (далее – автоматизированная система) (Правила, 2021). Однако, несмотря на эту структурированность, выявлено системное ограничение: доминирование ручных операций на стадии разработки. Установлено, что даже при функционировании автоматизированной системы, критически важные задачи (экспертиза, корректировка) выполняются с использованием текстовых редакторов и локальных носителей. Эта практика выступает барьером для полной цифровизации, ограничивает масштабируемость

и препятствует внедрению современных интеллектуальных инструментов.

Анализ технической спецификации, регламентирующей оказание услуг по разработке и формированию базы тестовых заданий, а также по сопровождению мероприятий, связанных с проведением внешних экзаменов. Данный документ отражает формализованную, строго этапированную модель жизненного цикла тестового задания, включающую подготовку экспертов, разработку заданий, проведение двух независимых экспертиз, корректировку, апробацию и внедрение в автоматизированную систему (Договор №4/ГЗ, 2025).

Результаты анализа позволяют констатировать наличие ряда значимых достижений в области цифровизации оценочных процедур. В частности, реализованы:

- апробация тестовых заданий в электронном формате;

- внедрение интеллектуальных систем видеонаблюдения (прокторинга) для обеспечения прозрачности процедуры тестирования;

- централизация обработки результатов с последующей выдачей электронных сертификатов;

- автоматизированное сопровождение процесса присуждения образовательных грантов.

Эти инициативы свидетельствуют о стремлении к повышению эффективности и прозрачности процедур внешней оценки и соответствуют современным трендам в сфере образовательных измерений. Вместе с тем, выявлены и существенные ограничения, препятствующие полноценной цифровой трансформации системы.

Учитывая масштаб охвата – свыше 1 000 000 тестируемых и функционирование 43 региональных центров тестирования – можно утверждать о высокой степени институциональной устойчивости и логистической готовности системы внешней оценки к проведению массового тестирования. В то же время выявленные ограничения подчёркивают необходимость цифровой трансформации всех этапов жизненного цикла создания тестовых заданий: от автоматизированного конструирования на основе когнитивных моделей до внедрения систем аналитической обратной связи. Комплексная модернизация в данном направлении позволит значительно повысить эффективность, прозрачность и качество национальной платформы оценивания.

Анализ внутренних отчётов и материалов служебного документооборота – таких как приказы, служебные записки, программы

обучения, списки привлечённых специалистов и итоги аprobации – позволил уточнить фактическую сторону реализации процесса формирования базы тестовых заданий. Эти документы преимущественно содержат количественные сведения и управленические решения, но не детализируют этапы разработки заданий. Тем не менее, они отражают ключевые аспекты организационной логистики, критерии отбора и подготовки участников, а также последовательность реализуемых ме-

роприятий. Полученные данные дополняют нормативную картину, позволяя более полно понять структуру операционных процессов и отдельные вызовы, с которыми сталкиваются исполнители на практике.

Документооборот осуществлялся с использованием системы электронного делопроизводства Workspace. Все материалы были получены из внутреннего архива организации и охватывают период с января по август 2025 года. Обобщённые данные представлены в таблице 1.

Таблица 1 – Категории и количество внутренних документов, включённых в анализ

№	Тип документа	Кол-во	Краткое содержание / Характер документации
1	Служебные записки	18	Вопросы привлечения разработчиков, экспертов, распределение функций, согласование сроков
2	Приказы	12	Утверждение рабочих групп, графиков мероприятий, назначение ответственных исполнителей
3	Внутренние отчёты	7	Промежуточные итоги разработки заданий, результаты экспертизы и аprobации
4	Программы повышения квалификации	3	Содержание курсов, перечень тем, количество обученных участников
5	Списки привлечённых специалистов	5	Региональное распределение, состав команд по разработке и экспертизе
6	Материалы аprobации тестовых заданий	4	Количество участников аprobации, форма проведения, краткие аналитические выводы

Общий объём проанализированных документов составил 49 единиц.

Таким образом, результаты анализа позволяют констатировать наличие институциональной базы, инфраструктурной готовности и отдельных цифровых решений в рамках системы внешней оценки. Вместе с тем, выявленные ограничения указывают на необходимость трансформации всего жизненного цикла тестового задания: от автоматизированного проектирования с использованием когнитивных моделей до внедрения систем аналитической обратной связи. Реализация этих шагов позволит повысить эффективность, устойчивость и прозрачность национальной платформы оценки.

Важно отметить, что общая картина ограничений и необходимости трансформации, выявленная в ходе анализа документов, подтверждается и детализируется в свободных комментариях участников опроса (B22). Например, один из участников отметил: «Отсутствие единой платформы для всех этапов разработки за-

даний усложняет процесс и увеличивает время на согласование». Другие подчеркивали: «Необходимость доработки алгоритмов генерации заданий под специфику каждого предмета» и «Сопротивление преподавателей новым технологиям из-за недостаточного обучения».

Полученные в ходе анализа документации выводы находят прямое подтверждение в эмпирических данных, представленных в исследовании, проведённом авторами ранее и опубликованном (Алтыбаева Ш., 2025). В рамках данного исследования оценивалась эффективность автоматизированной системы управления тестовыми заданиями в реальных условиях. Результаты показали, что функциональность управления заданиями (классификация, экспорт) получила высокую оценку экспертов ($M=4,67$), что подтверждает эффективность отдельных цифровых инструментов. Однако ключевые показатели производительности и надёжности системы были оценены ниже ($M = 3,92$).

Комплексный анализ нормативных и эмпирических данных позволяет сделать вывод, что основным «узким местом» является не отсутствие автоматизации как таковой, а инфраструктурные ограничения существующей платформы. Полученные данные подтверждают, что текущий уровень цифровизации ограничен преимущественно автоматизацией управления уже созданными заданиями, и требует в то время как стадия разработки остаётся в значительной степени ручной. Это снижает гибкость системы и затрудняет её масштабируемость, особенно в условиях растущих требований к прозрачности, уникальности и безопасности национальных тестов.

Результаты опроса. С целью расширения эмпирической базы исследования было проведён онлайн-опрос среди представителей казахстанских высших учебных заведений, участвующих в процессе разработки, согласования и анализа экзаменационных материалов. Целью опроса было выявление текущих институциональных практик, уровня цифровизации, потребностей в автоматизации, а также барьеров, препятствующих внедрению национальной цифровой платформы управления жизненным циклом экзаменационных заданий.

Анкета охватывала шесть тематических блоков и включала как закрытые, так и шкалированные и открытые вопросы. Респонденты представляли широкий спектр должностных позиций – от преподавателей и методистов до сотрудников ИТ-отделов и административного управления, что обеспечило комплексное представление о текущем состоянии процессов в разных типах вузов (государственных и частных, педагогических, технических, гуманитарных и др.). Приглашения к участию были разосланы через систему Workspace, с приложенной ссылкой на онлайн-анкету. Уровень отклика составил 68,54 %, что является высоким показателем для целевых экспертных опросов и позволяет минимизировать смещение, обусловленное неучастием (non-response bias).

Общее количество респондентов составило 206, при этом анкета охватывала более чем 22 вузов (Рисунок 1) из различных регионов Казахстана. Такая выборка позволила не только оценить текущую институциональную готовность к цифровой трансформации, но и выявить закономерности, характерные для различных категорий учебных заведений.

Полученные данные были обработаны с применением методов описательной статистики (ча-

стотный анализ, средние значения, стандартное отклонение), а также корреляционного анализа для выявления взаимосвязей между ключевыми показателями.

Таблица 2 – Распределение респондентов по типу учебного заведения

Категория	%
Государственный ВУЗ	85.4%
Частный ВУЗ	4.9%
НАО/Акционерное общество	9.7%

Из 302 специалистов, вовлечённых в процесс формирования базы тестовых заданий, участие в опросе приняли 206 респондентов. Как показано в таблице 2, более 85% из них представляют государственные вузы, что подтверждает доминирующее положение данной категории учреждений в системе высшего образования Республики Казахстан. Такие данные указывают на необходимость фокусировки усилий по цифровой трансформации, прежде всего, в сегменте государственных образовательных организаций.

Анализ профиля учебных заведений (В2) показал преобладание технических (28,6%), гуманитарных (19,9%) и педагогических (17,5%) вузов, что соответствует их доминирующей роли в системе высшего образования. В опросе приняли участие преимущественно преподаватели ($\approx 83\%$), непосредственно вовлеченные в разработку экзаменационных заданий, а также заведующие кафедрами, административный персонал, методисты и ИТ-специалисты, что обеспечило всесторонний охват респондентов (В3), задействованных в экзаменационном процессе.

Анализ использования специализированных цифровых платформ (В10) показал, что около 65% респондентов активно применяют такие решения для разработки заданий. Ограничено использование отмечают 14% опрошенных, а 21% участников вовсе не используют специализированные платформы. Среди наиболее часто упоминаемых наименований платформ (В11) лидируют Platonus (39%) и Moodle (22%). Также были отмечены собственные университетские разработки и системы, такие как Socrative и Univer. Наиболее часто применяются (В5) Microsoft Word (75%). Насчёт соблюдения стандартизованных шаблонов (В7) порядка 75% респондентов указали, что стандарты полностью соблюдаются, ещё

21% – что они соблюдаются преимущественно. Только в единичных случаях отмечено их частичное применение или отсутствие.

Исследование этапов разработки экзаменационных материалов (В8) выявило, что наиболее распространёнными практиками являются формирование заданий преподавателями (78%), их последующее обсуждение на кафедрах (72%) и проведение внутренней рецензии (46%). При этом утверждение на уровне факультета и пилотирование применяются реже, а внешняя экспертиза используется лишь эпизодически.

Что касается сроков подготовки заданий (В9), то в большинстве случаев разработка материалов занимает от одной до четырёх недель (69%), тогда как 23% участников исследования сообщили о превышении этого срока более чем на четыре недели. Периоды подготовки менее одной недели встречаются крайне редко, составляя менее 6% от общего числа ответов.

Анализ использования специализированных цифровых платформ (В10) показал, что около 65% респондентов активно применяют такие решения для разработки заданий. Ограниченнное использование отмечают 14% опрошенных, а 21% участников вовсе не используют специализированные платформы. Среди наиболее часто упоминаемых наименований платформ (В11) лидируют Platonus (39%) и Moodle (22%). Также были отмечены собственные университетские разработки и системы, такие как Socrative и Univer.

Оценка уровня цифровизации (В12) процессов разработки материалов демонстрирует вы-

сокую степень автоматизации: 84% участников оценили этот показатель от 3 до 5 баллов по пятибалльной шкале, при этом треть из них поставили максимальную оценку.

В отношении мер информационной безопасности (В13) наиболее распространёнными практиками являются физическое ограничение доступа к материалам (43,7%), защита электронных ресурсов (35,4%) и паролирование файлов (22,3%). Менее часто фиксируется применение шифрования, логирования и двухфакторной аутентификации.

Систематическое использование таксономии Блума (В16) подтверждено 39% респондентов. Ещё 46% отметили частичный учёт данного подхода или его использование в качестве общей рекомендации. Десять процентов высших учебных заведений не применяют таксономию Блума.

Потребности в автоматизации (В18) сосредоточены преимущественно на создании и редактировании заданий (40%), формировании тестовых вариантов (40%) и автоматизированной проверке и анализе результатов (38–41%). Автоматизация проверки развёрнутых ответов и управление версиями материалов были указаны как менее приоритетные, однако сохраняют свою актуальность.

Частотный анализ, представленный на рисунке 1, отражает распределение ответов респондентов по ключевым аспектам цифровизации процессов разработки экзаменационных материалов в вузах.

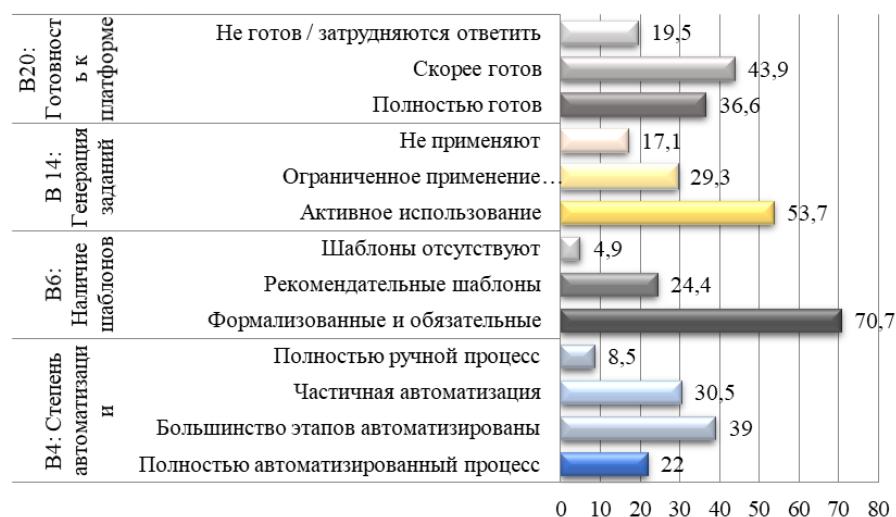


Рисунок 1 – Дескриптивная статистика (в %)

По результатам опроса, 61% вузов находятся на высоком уровне цифровизации, однако только 22% полностью автоматизировали разработку заданий, что свидетельствует о сохраняющемся потенциале для трансформации. Формализованные шаблоны применяют 70,7% вузов, тогда как их отсутствие зафиксировано лишь в 4,9% случаев, что подтверждает тенденцию к стандартизации.

Алгоритмы автоматической генерации используют 53,7% вузов, ещё 29,3% – частично, что указывает на растущий, но неравномерный

интерес к интеллектуальным решениям. Готовность к внедрению цифровой платформы выражали 80,5% респондентов; основными барьерами остаются технические, финансовые и организационные ограничения.

Проведённый корреляционный анализ (Таблица 3) с использованием коэффициента корреляции Спирмена позволил выявить значимые взаимосвязи между переменными, отражающими уровень цифровизации и характер её внедрения в вузах.

Таблица 3 – Корреляция Пирсона (числовые переменные)

№	Переменные	Коэффициент корреляции	p-значение	Интерпретация
1	Уровень автоматизации (B4) – Потребность в автоматизации (B17)	$r = -0.82$	< 0.01	Сильная отрицательная связь: при высокой автоматизации запрос на новые решения снижается.
2	Использование генерации заданий (B14) – Масштаб её применения (B15)	$r = 0.75$	< 0.05	Положительная связь: чем активнее использование, тем шире масштаб внедрения.

Установлена отрицательная корреляция между уровнем автоматизации экзаменационных процессов (B4) и потребностью в цифровизации (B17) ($\rho = -0.79$, $p < 0.01$), что свидетельствует о снижении запроса на новые решения при высоком уровне автоматизации.

Также выявлена положительная корреляция между использованием генерации заданий (B14) и масштабом её применения (B15) ($\rho = 0.72$, $p < 0.05$): активные пользователи чаще внедряют её на уровне всего вуза.

Таким образом, применение коэффициента Спирмена, обоснованное проверкой нормальности распределения, повышает валидность статистических выводов и количественно подтверждает связь между уровнем цифровой зрелости и институциональным спросом на автоматизацию.

Для анализа связи между готовностью вузов к внедрению цифровой платформы (B20) и барьерами (B21) использован критерий χ^2 , который показал статистически значимые ассоциации между уровнем готовности и основными ограничениями.

Таблица 4 – Результаты корреляционного анализа по критерию χ^2 между готовностью вузов и барьерами цифровизации

№	Переменные	хи-квадрат	p-значение	Интерпретация	Частота барьера среди «неготовых» вузов
3	Готовность к платформе (B20) – Барьеры внедрения (B21):				
	Техническая инфраструктура	15.3	< 0.001	Наиболее часто упоминаемый барьер среди вузов с низкой готовностью.	68%
	Финансовые ограничения	12.1	< 0.01	Ограниченнность бюджета существенно влияет на возможность внедрения.	54%
	Организационное сопротивление	9.8	< 0.01	Внутреннее сопротивление мешает цифровой трансформации.	41%

Результаты анализа (Таблица 4) продемонстрировали наличие статистически значимых ассоциаций по следующим направлениям:

- Недостаточная техническая инфраструктура ($\chi^2 = 15.3$, $p < 0.001$) – указана в 68% случаев среди вузов с низкой готовностью;
- Финансовые ограничения ($\chi^2 = 12.1$, $p < 0.01$) – 54% среди «неготовых» вузов;
- Организационное сопротивление изменениям ($\chi^2 = 9.8$, $p < 0.01$) – 41% респондентов отметили данный барьер.

Полученные результаты позволяют заключить, что восприятие цифровой трансформации в вузах существенно зависит не только от инфраструктурных и ресурсных факторов, но и от внутренней институциональной среды. Таким образом, барьеры внедрения платформы носят системный характер и требуют комплексного подхода при реализации цифровых решений в высшем образовании.

Триангуляция качественных и количественных данных показала, что основным узким местом в разработке тестовых заданий является отсутствие комплексной автоматизации на стадии создания, экспертизы и корректировки заданий, что проявляется в доминировании ручных операций и ограниченности существующей инфраструктуры для внедрения передовых интеллектуальных инструментов.

Исследование выявило значительный контраст между формализованными требованиями и фактической реализацией процесса разработки экзаменационных материалов, обнажая ключевые разрывы в цифровой трансформации. Несмотря на регламентированность и многоэтапность процесса, реальная практика по-прежнему сильно зависит от ручных операций, особенно на стадиях экспертизы и корректировки. Это создает фундаментальный барьер для масштабируемости системы и внедрения современных интеллектуальных инструментов, не предусмотренных текущим регламентом.

Несмотря на эти ограничения, в области цифровизации оценочных процедур достигнуты значительные успехи, включая электронную апробацию заданий, внедрение интеллектуальных систем прокторинга, централизованную обработку результатов и автоматизацию присуждения грантов. Масштаб охвата test-уровня (свыше 1 000 000 человек) и функционирование 43 региональных центров тестирования свидетельствуют о высокой институциональной устойчивости системы. Однако для полной цифровой

трансформации необходима модернизация всего жизненного цикла тестовых заданий – от автоматизированного конструирования до систем аналитической обратной связи.

Анализ внутренних отчётов дополняет картину, подтверждая, что, хотя некоторые цифровые инструменты эффективно используются для управления уже созданными заданиями ($M=4,67$), основным «узким местом» остается стадия разработки, которая в значительной степени продолжает быть ручной. Это связано не столько с полным отсутствием автоматизации, сколько с инфраструктурными ограничениями существующей платформы, что снижает гибкость и масштабируемость системы, особенно в условиях растущих требований к прозрачности и безопасности тестов.

Результаты онлайн-опроса среди представителей казахстанских вузов подтверждают эти выводы, детализируя текущие практики и потребности. Хотя большинство вузов (61%) демонстрируют высокий уровень цифровизации, лишь 22% полностью автоматизировали разработку заданий, указывая на нереализованный потенциал. Несмотря на высокую готовность к внедрению цифровой платформы (80,5%), существуют серьёзные барьеры. Корреляционный анализ выявил, что высокий уровень автоматизации снижает потребность в новых решениях, а анализ по критерию χ^2 чётко обозначил техническую инфраструктуру, финансовые ограничения и организационное сопротивление как основные препятствия, носящие системный характер и требующие комплексного подхода.

Методологические ограничения и перспективы. При интерпретации результатов следует учитывать методологические ограничения данного исследования. Выборка респондентов, хотя и репрезентативна для системы высшего образования Казахстана, может не в полной мере отражать позиции всех стейкхолдеров (например, представителей среднего профессионального образования, учащихся, работодателей). Кроме того, использование коэффициента Спирмена, хотя и корректно для порядковых данных, не позволяет делать выводы о линейной зависимости между переменными или устанавливать причинно-следственные связи. В дальнейших исследованиях рекомендуется: расширить выборку за счёт включения других категорий участников образовательного процесса; включить лонгитюдный компонент для анализа динамики процессов цифровизации.

Заключение

Данное исследование было посвящено анализу возможностей внедрения цифровой платформы для автоматизации разработки тестовых заданий в национальных экзаменационных системах Казахстана. Используя конвергентный смешанный дизайн с анализом документов и опросом 206 представителей вузов, мы выявили ключевые ограничения текущих процессов, институциональные потребности и потенциал цифровой трансформации.

Основные результаты показали, что процесс разработки заданий остается фрагментированным и зависимым от ручных операций, особенно на стадиях создания, экспертизы и корректировки. Несмотря на наличие автоматизированных систем (со средней оценкой $M=4,67$), отсутствие комплексной цифровизации ограничивает масштабируемость, валидность и безопасность тестов. Опрос выявил высокую готовность вузов к внедрению цифровой платформы (80,5%), однако технические, финансовые и организационные барьеры остаются значительными.

Успешное внедрение цифровой платформы зависит не только от технологической инфраструктуры, но и от институциональной готовности. Это включает предварительное обучение персонала, стандартизацию процессов тесто-разработки и реализацию поэтапного перехода. Такой подход способствует повышению прозрачности, управляемости и качества разработки заданий, обеспечивая соответствие международным стандартам, включая требования Болонского процесса. Внедрение единой цифровой платформы с функциями автоматической генерации заданий (AIG) позволяет преодолеть фрагментарность текущих практик, обеспечить стандартизацию, а также полную прослеживаемость всех этапов жизненного цикла экзаменационных материалов – от проектирования до анализа результатов. Для эффективного перехода рекомендуется внедрение стандартизованных метаданных на основе таксономии Блума, поэтапное развертывание платформы управления жизненным циклом заданий, интеграция с действующими информационными системами, поддержка двухязычных шаблонов и реализация встроенных психометрических модулей.

Литература

1. AQA. (2023). AQA acquires GradeMaker to transform exam development. <https://www.aqa.org.uk/news/aqa-acquires-grademaker>
2. American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. Washington, DC: AERA.
3. Brusilovsky, P., & Miller, P. (1999, October 24–30). Web-based testing for distance education. In P. De Bra & J. Leggett (Eds.), *Proceedings of WebNet '99, World Conference of the WWW and Internet* (pp. 149–154). Honolulu, HI: AACE.
4. Circi, R., Hicks, J., & Sikali, E. (2023). Automatic item generation: Foundations and machine learning-based approaches for assessments. *Frontiers in Education*, 8. <https://doi.org/10.3389/feduc.2023.858273>
5. Creswell, J. W., & Plano Clark, V. L. (2017). *Designing and conducting mixed methods research* (3rd ed.). Thousand Oaks, CA: SAGE Publications.
6. Gierl, M. J., Lai, H., & Fung, W. (2023). A suggestive approach for assessing item quality, usability and validity. *Perspectives on Medical Education*. <https://doi.org/10.1007/s40037-023-00835-0>
7. Grademarker. (n.d.). Bring your assessments into the digital age. <https://www.grademarker.com/>
8. taotesting.com. (n.d.). Setting the standard for excellence in digital assessment software. <https://www.taotesting.com>
9. unesco.org. (n.d.). Digital learning and transformation of education. <https://www.unesco.org/en/digital-education>
10. Williamson, D. M., Mislevy, R. J., & Bejar, I. I. (2006). Automated scoring of complex tasks in computer-based testing: An introduction to the special issue. *Journal of Technology, Learning, and Assessment*, 4(6). <https://ejournals.bc.edu/index.php/jtla/article/view/1655>
11. zakon.kz. (2021). ЕНТ-2021: как пройдет тестирование в электронном формате. <https://www.zakon.kz/stati/5061723-ent-2021-kak-proydet-testirovanie-v.html>
12. Алтыбаева, Ш. Б., & Скабаева, Т. С. (2025). Методические подходы к формированию тестовых заданий и цифровизации процедуры комплексного тестирования в компьютерном формате. *Педагогические измерения*, 1(1), 94–111. <https://testcenter-zhurnal.kz/index.php/sap/issue/view/1>
13. Договор №4/ГЗ от 12.03.2025. Министерство науки и высшего образования Республики Казахстан, Национальный центр тестирования. (2025).
14. Карданова, Е., & Иванова, А. (2023). Психометрические исследования: современные методы и новые возможности для образования. *Вопросы образования*, (3).

15. Концепция развития высшего образования и науки на 2023–2029 годы. (2023). <https://adilet.zan.kz/rus/docs/P2300000248>
16. Приказ Директора Национального центра тестирования от 4 июня 2021 года № 198-НҚ «Об утверждении Правил формирования базы тестовых заданий». (2021).
17. Сагиндиков, И. У., Динаева, Б. Б., & Убайдуллаева, Г. Ж. (2021). *Өлиеу және бағалау*. Нұр-Сұлтан.
18. Сагиндиков, И. У., Жумажанова, С. К., Ауезханова, А. Ж., & Тасболатова, М. Ш. (2023). *Жемілдіру және ынталандыру*. Астана.

References

- Altybayeva, Sh. B., & Skabayeva, T. S. (2025). Metodicheskiye podkhody k formirovaniyu testovykh zadaniy i tsifrovizatsii protsedury kompleksnogo testirovaniya v kompyuternom formate [Methodological approaches to the development of test items and the digitalization of the comprehensive computer-based testing procedure]. *Pedagogicheskiye izmereniya*, 1(1), 94–111. <https://testcenter-zhurnal.kz/index.php/sap/issue/view/1> (in Russian)
- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. Washington, DC: AERA.
- AQA. (2023). AQA acquires GradeMaker to transform exam development. <https://www.aqa.org.uk/news/aqa-acquires-grademaker>
- Brusilovsky, P., & Miller, P. (1999, October 24–30). Web-based testing for distance education. In P. De Bra & J. Leggett (Eds.), *Proceedings of WebNet'99, World Conference of the WWW and Internet* (pp. 149–154). Honolulu, HI: AACE.
- Circi, R., Hicks, J., & Sikali, E. (2023). Automatic item generation: Foundations and machine learning-based approaches for assessments. *Frontiers in Education*, 8. <https://doi.org/10.3389/feduc.2023.858273>
- Creswell, J. W., & Plano Clark, V. L. (2017). *Designing and conducting mixed methods research* (3rd ed.). Thousand Oaks, CA: SAGE Publications.
- Dogovor №4/GZ ot 12.03.2025. Ministerstvo nauki i vysshego obrazovaniya Respubliki Kazakhstan, Natsional'nyy tsentr testirovaniya [Contract No. 4/GZ dated 12.03.2025. Ministry of Science and Higher Education of the Republic of Kazakhstan, National Testing Center]. (2025). (in Russian)
- Gierl, M. J., Lai, H., & Fung, W. (2023). A suggestive approach for assessing item quality, usability and validity. *Perspectives on Medical Education*. <https://doi.org/10.1007/s40037-023-00835-0>
- Grademarker. (n.d.). Bring your assessments into the digital age. <https://www.grademarker.com/>
- Kardanova, E., & Ivanova, A. (2023). Psikhometricheskiye issledovaniya: sovremennyye metody i novyye vozmozhnosti dlya obrazovaniya [Psychometric studies: Modern methods and new opportunities for education]. *Voprosy obrazovaniya*, (3). (in Russian)
- Kontseptsiya razvitiya vysshego obrazovaniya i nauki na 2023–2029 gody [Concept for the development of higher education and science for 2023–2029]. (2023). <https://adilet.zan.kz/rus/docs/P2300000248> (in Russian)
- Prikaz Direktora Natsional'nogo tsentra testirovaniya ot 4 iyunya 2021 goda № 198-NQ «Ob utverzhdenii Pravil formirovaniya bazy testovykh zadaniy» [Order of the Director of the National Testing Center dated June 4, 2021, No. 198-NQ “On approval of the Rules for the formation of the test item bank”]. (2021). (in Russian)
- Sagindikov, I. U., Dinayeva, B. B., & Ubaydullayeva, G. Zh. (2021). *Ölsheu zhane bagalau* [Measurement and evaluation]. Nur-Sultan. (in Kazakh)
- Sagindikov, I. U., Zhumazhanova, S. K., Auezhanova, A. Zh., & Tasbolatova, M. Sh. (2023). *Zhetildiru zhane yntalandyrul* [Improvement and motivation]. Astana. (in Kazakh)
- taotesting.com. (n.d.). Setting the standard for excellence in digital assessment software. <https://www.taotesting.com>
- unesco.org. (n.d.). Digital learning and transformation of education. <https://www.unesco.org/en/digital-education>
- Williamson, D. M., Mislevy, R. J., & Bejar, I. I. (2006). Automated scoring of complex tasks in computer-based testing: An introduction to the special issue. *Journal of Technology, Learning, and Assessment*, 4(6). <https://ejournals.bc.edu/index.php/jtla/article/view/1655>
- zakon.kz. (2021). ENT-2021: kak proydet testirovaniye v elektronnom formate [UNT-2021: How testing will be conducted in electronic format]. <https://www.zakon.kz/stati/5061723-ent-2021-kak-proydet-testirovanie-v.html> (in Russian)

Сведения об авторах:

- Абдрасилов Болатбек Серикбаевич – академик НАН РК, директор, Национальный центр тестирования (Астана, Казахстан, e-mail: Bolatbek_s@mail.ru);
Алтыбаева Шугыла Болатовна – руководитель управления, магистр экономических наук, Национальный центр тестирования (Астана, Казахстан, e-mail: shugyla.altybayeva@gmail.com);
Ерсултанова Гульнур Серикбаевна – докторант, руководитель управления по развитию системы КАЗТЕСТ, Национальный центр тестирования (Астана, Казахстан, e-mail: ersultanovagulnur@gmail.com);
Жуматова Индира Жаналыковна – заведующая лабораторией, магистр гуманитарных наук, Национальный центр тестирования (Астана, Казахстан, e-mail: i.zh.86@bk.ru);
Туралбаева Кенжекул Турганбаевна – заведующая лабораторией, Национальный центр тестирования (Астана, Казахстан, e-mail: kengebahit@mail.ru).

Information about authors:

Abdrasilov Bolatbek Serikbayevich – Academician of the National Academy of Sciences of the Republic of Kazakhstan, Director, National Testing Center, (Astana, Kazakhstan, e-mail: Bolatbek_s@mail.ru)

Altybaeva Shugyla Bolatovna – Head of Department, Master of Economics, National Testing Center, (Astana, Kazakhstan, e-mail: shugyla.altybayeva@gmail.com)

Yersultanova Gulnur Serykbaevna (contact person) – doctoral student, Head of the Department for the Development of the KAZTEST System, National Testing Center (Astana, of Kazakhstan, e-mail: ersultanovagulnur@gmail.com)

Zhumatova Indira Zhanalykovna – head of the Laboratory, Master of Humanities, National Testing Center (Astana, Kazakhstan, e-mail: i.zh.86@bk.ru)

Turalbayeva Kenzhekul Turganbayevna – Head of laboratory, National Testing Center, (Astana, Kazakhstan, e-mail:kengebahit@mail.ru)

Авторлар туралы мәлімет:

Абдрасилов Болатбек Серикбаевич – КР ҮФА академигі, директор, Үлттүк тестілеу орталығы (Астана, Қазақстан,- e-mail:Bolatbek_s@mail.ru)

Алтыбаева Шүгыла Болатовна – басқарма басшысы, экономика ғылымдарының магистрі, Үлттүк тестілеу орталығы (Астана, Қазақстан, e-mail: shugyla.altybayeva@gmail.com)

Ерсултанова Гульнур Серикбаевна – докторант, ҚАЗТЕСТ жүйесін дамыту жөніндегі басқарма басшысы, Үлттүк тестілеу орталығы (Астана, Қазақстан,e-mail: ersultanovagulnur@gmail.com)

Жұматова Индира Жаналыковна – зертхана меншерушісі, гуманитарлық ғылымдар магистрі, Үлттүк тестілеу орталығы (Астана, Қазақстан, e-mail: i.zh.86@bk.ru)

Туралбаева Кенжекүл Тұрганбаевна – зертхана меншерушісі, Үлттүк тестілеу орталығы (Астана, Қазақстан, e-mail: kengebahit@mail.ru)

Поступила: 18.09.2025

Принята: 01.12.2025