

# ЦИФРОВЫЕ РЕШЕНИЯ ДЛЯ ПОВЫШЕНИЯ КАЧЕСТВА УПРАВЛЕНИЯ УНИВЕРСИТЕТОМ

---



НЕКОММЕРЧЕСКОЕ АКЦИОНЕРНОЕ ОБЩЕСТВО  
«КАЗАХСКИЙ НАЦИОНАЛЬНЫЙ УНИВЕРСИТЕТ ИМЕНИ АЛЬ-ФАРАБИ»

---

# ЦИФРОВЫЕ РЕШЕНИЯ ДЛЯ ПОВЫШЕНИЯ КАЧЕСТВА УПРАВЛЕНИЯ УНИВЕРСИТЕТОМ

*Монография*

*под общей редакцией  
К.Е. Смагулова и М.Е. Мансуровой*

Алматы  
«Smart University Press»  
2022

УДК 378(035.3)

ББК 74.58

Ц 75

*Рекомендовано к изданию Ученым советом  
Факультета информационных технологий КазНУ им. аль-Фараби  
(протокол № 1 от 13.09.2022)*

**Рецензенты:**

д.т.н., профессор **Б.К. Синчев**

доктор PhD **Б.С. Дарибаев**

Ц 75 **Цифровые** решения для повышения качества управления университетом: монография / Под общей редакцией К.Е. Смагулова и М.Е. Мансуровой. – Алматы: Smart University Press, 2022. – 133 с.

**ISBN 978-601-269-123-8**

Внедрение цифровых решений является приоритетом как государственных органов, так и частного сектора, субъектов бизнеса и организаций образования. В монографии представлены результаты исследований по разработке информационно-аналитических платформ и автоматизации процессов управления университетом, которые позволяют сократить время обработки данных, формировать профили пользователей, выработать стратегию дальнейшего развития, при необходимости калибровать деятельность для достижения поставленных целей. Первая глава монографии посвящена вопросам проектирования и разработки интеллектуальной информационно-аналитической системы с визуализацией результатов агрегации и анализа данных о показателях здоровья студентов. Во второй главе монографии приводится анализ существующих библиометрических методов оценки научной производительности ученых и организаций на основе публикационной активности в рейтинговых изданиях, индексируемых в международных базах данных. Представлена автоматизированная система по оценке публикационной активности авторов и организаций РК на основе данных из базы Scopus и с применением метода анализа с учетом специфических требований, стоящих перед казахстанскими учеными в настоящее время и соответствующих требованиям нормативно-правовых актов Республики. Дано техническое описание системы, а также методология ранжирования ученых и организаций Казахстана на основе формирования их рейтинга по количественным и качественным показателям публикационной активности. Монография предназначена научным работникам, специалистам в области управления университетом, преподавателям вузов, докторантам, магистрантам и студентам.

**УДК 378(035.3)**

**ББК 74.58**

ISBN 978-601-269-123-8

© КазНУ имени аль-Фараби, 2022

## **ВВЕДЕНИЕ**

В современном мире информационные технологии играют все более важную роль во многих аспектах деятельности человека, в том числе в системе высшего образования. Внедрение цифровых решений является приоритетом как государственных органов, так и частного сектора, субъектов бизнеса и организаций образования.

Ключевым аспектом исследований, представленных в первой главе монографии, является разработка цифрового профиля здоровья студента, преобразование необработанной информации, полученной из различных источников, в знания и рекомендации, поддерживающие процесс принятия решений. Цель организаций образования – помочь молодым людям принять более здоровые формы поведения на ранних этапах жизни и адаптироваться к изменениям, которые связаны с особенностями возраста и среды. Актуальность обсуждаемых вопросов определяет необходимость комплексных исследований, позволяющих оценивать текущее состояние студентов, и разрабатывать индивидуализированные подходы к формированию профилактической среды и укреплению здоровья в условиях образовательных учреждений. В рамках указанной научной проблемы решается целый комплекс задач, связанных с формализацией, структурированием, анализом разнообразных наборов данных, касающихся состояния здоровья студентов, поиском в исследуемых данных скрытых закономерностей, шаблонов и связей, агрегированием данных и их визуализацией. Разработанные в ходе исследования методы и алгоритмы интеллектуального анализа данных позволяют определить показатели, имеющие наибольшее влияние на состояние здоровья и благополучие молодежи, а также выявить факторы риска нарушений здоровья. Выполнение этих задач позволит разработать актуальную, но отсутствующую на сегодняшний день систему, осуществляющую на регулярной и непрерывной основе сбор и анализ информации о состоянии здоровья учащейся молодежи, имеющую большое значение с точки зрения устойчивого национального развития.

Во второй главе члены авторского коллектива приводят результаты исследования, направленного на разработку автоматизированной системы для ранжирования ученых и организаций Казахстана на основе показателей публикационной активности в базе данных Scopus, а также балловой системы на основе этих показателей. Авторы анализируют имеющиеся в научной литературе теории и библиометрические методы оценки научной производительности ученых и организаций на основе публикационной активности в рейтинговых изданиях, индексируемых в международных базах данных.

Для формирования авторской методологии оценки научной производительности исследователи приводят анализ действующих в различных странах методологий и систем оценки эффективности деятельности ученых и университетов на основе количественных и качественных показателей, оценивающих вклад ученого, организации, страны в написание научной статьи, цитируемость публикаций, международная коллаборация и др.

На основе проведенного анализа литературы и опыта зарубежных стран приведена техническая характеристика разработанной авторской автоматизированной системе по оценке публикационной активности авторов и организаций Республики Казахстан на основе данных из базы Scopus, выгружаемых посредством самовыполняемого программного приложения по API-каналу. Важность данной системы заключается в том, что авторы разработали собственные индикаторы и параметры, соответствующие требованиям к публикациям в международных журналах, отраженных в нормативно-правовых актах Казахстана. Так, например, в системе был внедрен фильтр по отображению только тех статей, которые опубликованы в журналах с показателем CiteScore выше 25 по базе данных Scopus, что является пороговым требованием к журналам, предъявляемым для опубликования статей PhD-докторантами согласно Правилам присуждения степеней.

Таким образом, представленные в монографии цифровые решения будут способствовать повышению качества управления университетом, внесут вклад в дальнейшую цифровизацию сферы высшего образования.

# **1. ИНТЕЛЛЕКТУАЛЬНАЯ ИНФОРМАЦИОННО-АНАЛИТИЧЕСКАЯ СИСТЕМА «ЦИФРОВОЙ ПРОФИЛЬ ЗДОРОВЬЯ СТУДЕНТА»**

*Тюлепбердинова Г.А., Мансурова М.Е., Сарсембаева Т.С.,  
Сарсенова Л.К., Амирханова Г.А.*

В данной главе представлены результаты исследований по разработке и внедрению информационно-аналитической системы оценки состояния здоровья студентов для формирования комплекса мероприятий по профилактике заболеваний и улучшению качества жизни молодежи с применением методов интеллектуального анализа данных. В главе приводится анализ существующих решений по оценке состояния здоровья молодого поколения. Представлена архитектура авторской интеллектуальной информационно-аналитической системы «Цифровой профиль здоровья студента», которая может быть интегрирована в систему управления учебным процессом высшего учебного заведения.

Предлагаемые исследования призваны ответить на актуальный вызов, связанный с недостаточным уровнем осведомленности населения и государственных административных органов о здоровье молодого поколения РК. На цифровой платформе университета развернуты и предоставлены заинтересованным лицам интеллектуальные инструменты поддержки принятия решений на основе сбора и анализа данных о состоянии здоровья студентов. Ключевым аспектом исследований является разработка цифрового профиля здоровья студента, преобразование необработанной информации, полученной из различных источников, в знания и рекомендации, поддерживающие процесс принятия решений. В рамках указанной научной проблемы решается целый комплекс задач, связанных с формализацией, структурированием, анализом разнообразных наборов данных, касающихся состояния здоровья студентов, поиском в исследуемых данных

скрытых закономерностей, шаблонов и связей, агрегированием данных и их визуализацией. Разработанные в ходе исследования методы и алгоритмы интеллектуального анализа данных позволяют определить показатели, имеющие наибольшее влияние на состояние здоровья и благополучие молодежи, а также выявить факторы риска нарушений здоровья. Представленные результаты получены в рамках совместных междисциплинарных исследований специалистов в области информационных технологий и здравоохранения.

*Данная глава монографии подготовлена в рамках выполнения научного проекта на тему АР09260767 «Разработка интеллектуальной информационно-аналитической системы оценки состояния здоровья студентов Казахстана» грантового финансирования Комитета науки МОН РК на 2021-2023 гг.*

### **1.1. Анализ существующих решений по оценке состояния здоровья молодого поколения**

Студенты сталкиваются с различными потенциально серьезными рисками для здоровья во время учебы в университете, и их образ жизни может иметь прямое влияние на здоровье и поведение в отношении здоровья в более позднем возрасте<sup>1 2 3</sup>. В то же время студенты университетов – это студенты поколения Z, которые имеют легкий доступ к Интернету и новым технологиям<sup>4</sup>. Применение инструментов цифрового здравоохранения и вовлечение студентов университетов в исследования состояния здоровья молодых людей позволит выработать стратегию университе-

---

<sup>1</sup> “Health behavior in school-aged children”, MRC/CSO Social and Public Health Sciences Unit, University of Glasgow [Электронный ресурс]. – URL: <https://hbsc.org/>

<sup>2</sup> Нефедовская Л.В., Состояние и проблемы здоровья студенческой молодежи // Литтерра, Москва. – 2017. – стр.192.

<sup>3</sup> Баранова А.А., Кучмы В.Р., Физическое развитие детей и подростков Российской Федерации: Сб.мат-лов (выпуск VI) //ПедиатрЪ, Москва – 2013. – стр. 192.

<sup>4</sup> “Understanding Generation Z in the workplace”, 2022 Global Automotive Consumer Study [Электронный ресурс]. – URL: <https://www2.deloitte.com/us/en/pages/consumer-business/articles/understanding-generation-z-in-the-workplace.html>

та по поддержке студентов. Вмешательства в области цифрового здравоохранения открывают многообещающие возможности для укрепления здоровья, профилактики заболеваний и оказания помощи этой конкретной группе населения. Цифровые ресурсы очень доступны и при активном использовании могут быть ценным инструментом, побуждающим студентов задуматься о своем образе жизни.

### ***1.1.1. Проблемы оценки состояния здоровья молодежи в мире и РК***

По данным Комитета по статистике Министерства национальной экономики Республики Казахстан за 2018 год, количество молодых людей в возрасте от 14 до 29 лет составило около 21,7% от общей численности населения, то есть 3 900 834 человека<sup>5</sup>, которая представляет собой самую маленькую демографическую когорту. Причина столь низкого показателя – «демографическая яма 90-х», когда в 1995-2005 гг. рождаемость стала катастрофически низкой<sup>6 7</sup> (Рис. 1.1). Также стоит отметить, что этому процессу способствует миграционный отток молодежи.

В результате социально-экономических реорганизаций на протяжении последних 20-25 лет в Казахстане исследуются проблемы в области молодежной популяции. Отмечается уменьшение количества молодежи, вследствие «демографической ямы 90-х» и ухудшение состояния их здоровья. Данные проблемы определяют необходимость комплексных исследований, позволяющих на доклиническом уровне оценивать текущее состояние

---

<sup>5</sup> Абдрахманова Ш.З., Ахметов В.И., Адаева А.А., Слажнева Т.И., Факторы, формирующие здоровье и детей и подростков Казахстана. Health behavior in school-aged children 2018 // *Национальный отчет и Национальный центр общественного здравоохранения, Нур-Султан.* – 2019. – стр. 150.

<sup>6</sup> Қалиев Т.Б., Қайдарова Ә.С., Қаримова Ж.К., Әшімханова Д.Ә., Маульшариф М.М., Насимова Г.О., Негай Н.А., Сыдықназаров М. Қ., Шаповал Ю.В., Мусатаева Ф.М., Оқасова Г.Е., Қасымбеков А.М., Молодежь Казахстана – 2018 и Youth of Kazakhstan – 2018//Национальный доклад и National report, Астана. – 2018. – стр. 410.

<sup>7</sup> «Исследование распространенности, глубинных причин и факторов риска и защиты в области суицида и суицидальных попыток в Республике Казахстан» // Детский Фонд ООН (ЮНИСЕФ) в Республике Казахстан, Астана. – 2014. – стр. 108.



студентов, и разрабатывать индивидуализированные подходы к формированию профилактической среды и укреплению здоровья в условиях образовательных учреждений.

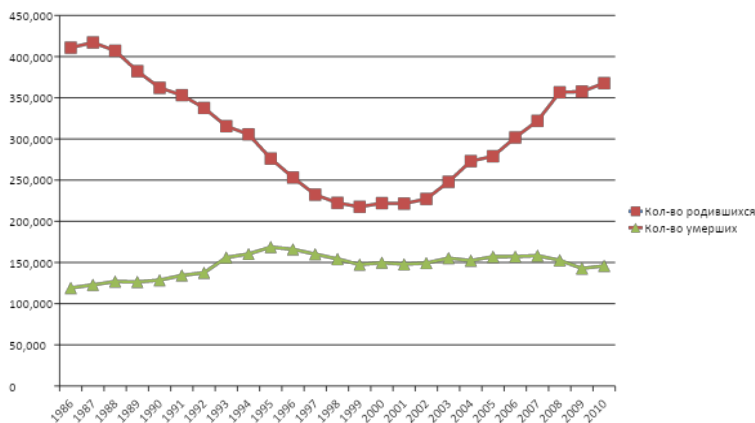


Рисунок 1.1 - Демографическая статистика РК за 1986-2010 гг.

Студенты университетов составляют почти две трети всей молодежи в странах Организации экономического сотрудничества и развития<sup>8</sup> (ОЭСР). Их здоровье и благополучие как потенциальных будущих лидеров, политиков и менеджеров являются глобальным приоритетом общественного здравоохранения. Хотя их можно рассматривать как привилегированное здоровое население, студенты университетов часто сообщают о плохом здоровье. У них относительно высокий уровень заболеваний, передающихся половым путем, и воспалительных заболеваний из-за рискованных сексуальных практик<sup>9</sup>; они подвержены риску хронических заболеваний из-за малоподвижного поведения и часто сообщают о проблемах психиче-

<sup>8</sup> «Организация экономического сотрудничества и развития (ОЭСР). Краткий обзор образования: показатели ОЭСР» // ОЭСР, Париж – 2017.

<sup>9</sup> Lechner K.E., Garcia C.M., Frerich E.A., Lust K., Eisenberg M.E., College students' sexual health: personal responsibility or the responsibility of the college? [Электронный ресурс]. – 2013. – URL: <http://europepmc.org/abstract/MED/23305542>

ского здоровья, таких как стресс, беспокойство или депрессия, которые часто возникают из-за академической нагрузки и тоски по дому<sup>10</sup>, проблемного употребления алкоголя<sup>11</sup> и наркотиков.

Важность здоровья подростков для мирового сообщества под лозунгом «Глобальный приоритет – здоровая молодежь» подчеркивается в стратегических документах международных медицинских организаций UNICEF, WHO, SHE (Schools for Health in Europe), HBSC (Health behavior in school-aged children), YHO (Youth Health organization), UNFPA (фонд ООН в области народонаселения), в стратегии Health 2020<sup>12 13 14 15</sup>.

В РК также здоровье молодежи как глобальная повестка отражена в Законе РК «О государственной молодежной политике»<sup>16</sup>, в Кодексе РК «О здоровье народа и системе здравоохранения»<sup>17</sup>, в Концепции «Государственной программы улучшения здоровья

<sup>10</sup> Hunt J., Eisenberg D., Mental health problems and help-seeking behavior among college students – 2010. – URL: <https://pubmed.ncbi.nlm.nih.gov/20123251/>

<sup>11</sup> Quinn P.D., Fromme K., Alcohol use and related problems among college students and their noncollege peers: the competing roles of personality and peer influence [Электронный ресурс]. – 2011. – URL: <https://europepmc.org/article/MED/21683044>

<sup>12</sup> «Основы европейской политики и стратегия для XXI века. Здоровье-2020» [Электронный ресурс]. – 2020. – URL: <http://www.euro.who.int/>

<sup>13</sup> Patton G.C., Sawyer S.M., Santelli J.S., Ross D.A., Afifi R., Allen N.B., Arora M., Azzopardi P., Baldwin W., Bonell C., Kakuma R., Kennedy E., Mahon J., McGovern T., Ali H. Mokdad, Patel V., Petroni S., Reavley N., Taiwo K., Waldfogel J., Wickremarathne D., Barroso C., Bhutta Z., Fatusi O.A., Mattoo A., Diers J., Fang J., Ferguson J., Ssewamala F., Viner R.M., A Lancet commission on adolescent health and wellbeing [Электронный ресурс]. – 2016. – URL: <https://pubmed.ncbi.nlm.nih.gov/27174304/>

<sup>14</sup> Currie C., Социальные детерминанты здоровья и благополучия подростков. Исследование «Поведение детей школьного возраста в отношении здоровья»: международный отчет по результатам обследования 2009–2010 гг. //Европейское региональное бюро ВОЗ Серия «Политика охраны здоровья детей и подростков» – 2012.

<sup>15</sup> Barrett M.A., Humblet O., Hiatt R.A., Adler N.E., Big Data and Disease Prevention: From Quantified Self to Quantified Communities [Электронный ресурс]. – 2013. – URL: <https://doi.org/10.1089/big.2013.0027>

<sup>16</sup> Закон Республики Казахстан «О государственной молодежной политике» от 9 февраля 2015 года № 285-V ЗРК [Электронный ресурс]. – URL: <https://doi.org/10.1089/big.2013.0027> <http://edu.resurs.kz/elelegal/molodezhnaya-politika>

<sup>17</sup> О здоровье народа и системе здравоохранения. Кодекс Республики Казахстан от 7 июля 2020 года № 360-VI ЗРК [Электронный ресурс]. – URL: <https://adilet.zan.kz/rus/docs/K2000000360v>

населения на 2020-2025 гг.», образован «Совет по молодежной политике» при Президенте РК<sup>18</sup>.

### ***1.1.2. Цифровые технологии для оценки состояния здоровья молодежи***

«Центр США по контролю и профилактике заболеваний» (CDC – Centers for Disease Control and Prevention) занимается защитой и безопасностью общественного здоровья, которая осуществляется посредством контроля и предотвращения заболеваний, травм и инвалидности<sup>19 20</sup>. Система наблюдения за рискованным поведением молодежи YRBSS (The Youth Risk Behavior Surveillance System) разработана CDC. Система отслеживает поведение, которое может быть связано с риском для здоровья среди молодых людей на национальном, государственном, территориальном и местном уровнях<sup>21</sup>. Более того, с помощью YRBSS Youth Online<sup>22</sup> можно фильтровать и сортировать данные по различным параметрам, создавать собственные таблицы, карты, диаграммы, а также выполнять статистические тесты по местоположению и состоянию здоровья. В онлайн-системе доступны инструменты анализа и визуализации данных как для отдельных регионов, так и для страны в целом, а также портал данных с необходимыми соответствующими наборами данных и информационными панелями. Для анализа данных в системе YRBSS используются такие статистические программные пакеты, как SUDAAN, SAS, Stata, SPSS и Epi Info, а также пакет анализа языка программирования R<sup>23</sup>. Собранные, обработанные и проанализированные данные позволяют нам получить оценки

---

<sup>18</sup> Указ Главы государства от 1 июля 2008 г. № 625

<sup>19</sup> Catherine N.R., Georgianne F.T., Laura K., Tim M., Shannon L.M., Caitlin L.M., Sarah M.L., Michelle K.B., Frances A., Kathleen A.E.

<sup>20</sup> Catherine N.R., Georgianne F.T., Laura K., Tim M., Shannon L.M., Caitlin L.M., Sarah M.L., Michelle K.B., Frances A., Kathleen A.E., Youth Risk Behavior Surveillance System . [Электронный ресурс]. – 2015. – URL: <https://www.cdc.gov/healthyschools/shi/index.htm>

<sup>21</sup> [Электронный ресурс]. – 2015. – URL: <https://www.cdc.gov/healthyyouth/data/yrbs/index.htm>

<sup>22</sup> [Электронный ресурс]. – 2015. – URL: <https://nccd.cdc.gov/YouthOnline/>

<sup>23</sup> [Электронный ресурс]. – 2015. – URL: <https://www.cdc.gov/healthyschools/about>

и тенденции риска для здоровья в поведении молодых людей; понимать ключевые факторы, влияющие на степень риска; и, кроме того, использовать эти данные для улучшения здоровья молодежи с помощью эффективных школьных и социальных программ, политики и подходов<sup>24</sup>.

Цифровые технологии открывают новые возможности для изучения показателей здоровья учащейся молодежи. Авторы подчеркивают важность проведения наблюдений за здоровьем и состоянием молодежи<sup>25</sup>. Они описывают платформу для обмена знаниями для поддержки уровня осведомленности о здоровье молодежи нескольких провинций Канады, по разработке политики, планировании и оценке программ для улучшения здоровья молодежи и профилактики хронических заболеваний.

HBSC (Health Behavior in School-age Children) является примером самого масштабного межнационального долговременного исследования здоровья детей и подростков. Последний отчет по результатам исследования за 2006-2018 гг. показал ряд положительных и отрицательных трендов<sup>26</sup>. Такого рода исследования служат основой для принятия решений как на национальном, так и на общемировом уровне, что, учитывая долю молодых людей в структуре популяции и их будущую социальную роль, имеет стратегическое значение. Особое место в мониторинге здоровья молодежи занимают вопросы психического благополучия. В инструментарии для школ и колледжей Англии представлен обзор апробированных и доказавших свою эффективность инструментов для измерения и мониторинга психического благополучия детей и молодежи<sup>27</sup>. В нем приведены примеры успешного

<sup>24</sup> [Электронный ресурс]. – 2015. – URL: <https://www.cdc.gov/healthyschools/about>

<sup>25</sup> Murnaghan D., Morrison W., Griffith E.J., Bell B.L., Duffley L.A., McGarry K., Manske S., Knowledge exchange systems for youth health and chronic disease prevention: a tri-provincial case study // *Chronic Diseases and Injuries in Canada* [Электронный ресурс]. – 2013. – URL: <https://pubmed.ncbi.nlm.nih.gov/23987222/>

<sup>26</sup> Heinz A., Claire van D., Kern R.M., Catunda C., Helmut erich W., Trends from 2006 - 2018 in Health Behaviour, Health Outcomes and Social Context of Adolescents in Luxembourg [Электронный ресурс]. – 2020. – URL: <https://www.researchgate.net/publication/342630277>.

<sup>27</sup> “Measuring and monitoring children and young people’s mental wellbeing”. *Toolkit for schools and colleges*. – 2006. – Vol.91(2).

применения онлайн инструментов, например, Warwick-Edinburgh Mental Wellbeing scale<sup>28</sup> и The Boxall Profile Online<sup>29</sup>.

Большое количество накопленных медицинских данных, хранящихся в цифровом виде, позволяют разрабатывать системы наблюдения и поддержки принятия ориентированных на пациента решений. Рекомендательные системы в здравоохранении позволяют заинтересованным лицам: пациентам, службам социальной и психологической поддержки, получать актуальную информацию и рекомендации<sup>30 31 32 33</sup>. Автоматизировать процессы принятия решений в области здравоохранения можно с помощью применения методов машинного обучения. В ходе выполнения проекта будет разработан рекомендательный функционал, который позволит принимать более информированное решение о том, как поддержать и улучшить состояние здоровья молодых людей.

### ***1.1.3. Применение опросников для сбора медицинских данных***

Многие поведенческие и связанные со здоровьем ситуации формируются в подростковом возрасте. Отношение молодых людей к здоровью и поведению/здоровью может оказать значительное влияние на здоровье во взрослом возрасте. Согласно статистике, в Казахстане выросли ключевые показатели рисков для здоровья подростков<sup>34</sup>. В настоящее время у более половины подростков диагностировано какое-либо хроническое заболевание. В этой связи проблема сохранения и укрепления здоровья молодежи

---

<sup>28</sup> [Электронный ресурс]. – URL: <https://warwick.ac.uk/fac/sci/med/research/platform/wemwbs/about/>

<sup>29</sup> [Электронный ресурс]. – URL: <https://new.boxallprofile.org/#how>

<sup>30</sup> Benjamin S., Constanze K., Mert A., Karim E.A., Literature Review on Medicine Recommender Systems // *International Journal of Advanced Computer Science and Applications*. – 2019. – P. 82.

<sup>31</sup> Qiwei H., Mengxin J., Inigo Martinez R. T., Manas G., Leid Z., A Hybrid Recommender System for Patient [Электронный ресурс]. – 2018. – URL: <https://www.researchgate.net/publication/330880005>.

<sup>32</sup> Joris F., Amine A.Y., Frédéric M.B.H., Brigitte D., Alpha D., Visual instance-based recommendation system for medical data mining // *Procedia Computer Science* 112. – 2017. – P. 1747-1754.

<sup>33</sup> Martin W., Daniel P., Concepts, Requirements, Technical Basics and Challenges // *International Journal of Environmental Research and Public Health* – 2014. – P. 2580 – 2607.

<sup>34</sup> «Анализа положения в области народонаселения в Республике Казахстан» [Электронный ресурс]. – 2019. – URL: <https://kazakhstan.unfpa.org>

жи имеет высокий социальный и общественный приоритет для страны.

Медицинские опросники являются одним из действенных инструментов сбора информации. Они разрабатываются компетентными специалистами, проходят стандартизацию и валидацию. Для оценки состояния детей и подростков в международной практике используется межнациональное исследование HBSC, разработанное Всемирной организацией здравоохранения<sup>35</sup>. Исследовательская сеть HBSC объединяет людей с широкими знаниями в таких областях, как клиническая медицина, эпидемиология, биология человека, педиатрия, педагогика, психология, здравоохранение, государственная политика и социология.

Целью исследования HBSC является использование собранных данных:

- для формирования нового отношения к здоровью и благополучию молодого поколения;
- для получения более глубокого представления о социальных детерминантах здоровья;
- для проведения мероприятий с целью улучшения жизни подростков.

Подростковый период является периодом, который оказывает большое влияние на развитие человека, поскольку поведение, которое формируется в этот период, может продолжаться даже при переходе во взрослую фазу, что влияет на такие факторы, как психическое здоровье, развитие жалоб на здоровье, курение, диета, уровень физической активности и потребление алкоголя. Чтобы разработать эффективную политику укрепления здоровья, направленную на улучшение здоровья и снижение неравенства в области здоровья среди молодежи, необходимо признать особые биологические, эмоциональные и социальные изменения, происходящие в этот период жизненного цикла, и связанные с ними проблемы. На здоровье и благосостояние подростков оказывают сильное влияние социальные факторы, непосредственно связанные с окружающей средой молодежи, включая семью, школу и общество.

---

<sup>35</sup> [Электронный ресурс]. – URL: <https://hbhc.org/>

Для всех исследований HBSC разрабатывается стандартизированный протокол исследования, обеспечивающий теоретическую основу тем исследований и процедур сбора и анализа данных. Протокол направлен на обеспечение безопасности сопоставимых данных. Члены исследовательской сети HBSC сотрудничают в подготовке этого международного протокола исследований для каждого четырехлетнего исследования. Протокол исследования содержит подробную информацию и рекомендации, включающие: концептуальные рамки исследования; научные обоснования для каждой тематической области исследования; международную стандартную версию опросников и инструкций по эксплуатации (например, рекомендуемый формат, порядок вопросов и инструкцию по переводу); всеобъемлющее руководство по методологии исследования, включая рекомендации по отбору национальных наборов данных для экспорта в Международный банк данных, процедуры сбора данных и подготовки; и положения, связанные с использованием данных HBSC и международной публикацией<sup>36</sup>.

Международный стандартный опросник состоит из трех уровней вопросов, используемых для создания национальных инструментов исследования: основные вопросы, которые должна охватывать каждая страна для создания международного набора данных; дополнительные подборки вопросов по конкретным тематическим областям, которые страны могут выбрать; и вопросы по конкретным странам, связанные с вопросами национального значения. Вопросы включают ряд показателей здоровья и поведения, связанных со здоровьем, а также жизненные ситуации молодых людей. Международная сеть организована вокруг ряда целевых и тематических групп, связанных между собой различными направлениями:

- Социальный контекст: общение с родителями, друзьями и одноклассниками, университетская среда, социально-экономический показатель семьи;

---

<sup>36</sup> [Электронный ресурс]. – URL: [https://www.who.int/europe/initiatives/health-behaviour-in-school-aged-children-\(hbsc\)-study](https://www.who.int/europe/initiatives/health-behaviour-in-school-aged-children-(hbsc)-study)

- Показатели здоровья: оценка здоровья студентов, представления об их организме, уровень удовлетворенности жизнью, жалобы на здоровье, травмы, масса тела;

- Поведение, способствующее здоровью: пищевое поведение и диета, гигиена полости рта, физическая активность, подвижный и малоподвижный образ жизни, ожирение;

- Поведение, угрожающее здоровью: курение, употребление алкоголя, употребление наркотиков, секс и методы защиты, участие в драках, запугивание.

Результаты HBSC показывают, как здоровье молодых людей меняется с детства и подросткового возраста, когда они переходят во взрослую жизнь. Собранные данные показывают, какие меры следует предпринять, чтобы контролировать здоровье молодежи, понимать социальные детерминанты здоровья и, что еще более важно, помочь детям и подросткам сформировать позитивное поведение в отношении здоровья, которое может сохраняться в течение всей жизни<sup>37</sup>.

Многие аспекты поведения, связанные со здоровьем и привычками, которые приводят к болезням (сердечно-сосудистые заболевания, диабет, рак, проблемы с психическим здоровьем), инвалидность у взрослых и преждевременная смерть возникают в подростковом возрасте. Исследование HBSC помогает понять, почему это происходит.

Известно, что на состояние здоровья молодежи влияет сложный комплекс различных факторов, среди которых особое место занимает учебная нагрузка, связанная со стрессовыми ситуациями, осложненными интенсивной умственной деятельностью, нарушением режима труда и отдыха, физической активностью, вредными привычками. В связи с этим студенты относятся к отдельной группе риска как социальная группа, подверженная риску развития многих заболеваний. Особенностью оценки состояния здоровья в молодом возрасте является предварительное

---

<sup>37</sup> Jason M. N., New Findings From the Health Behaviour in School-aged Children (HBSC) Survey: Social Media, Social Determinants, and Mental Health. Adolescent Health. – 2020. - Vol. 66, No 6.



определение возможных заболеваний и поведенческих индикаторов здоровья в зрелом возрасте и проведение мероприятий на необходимом для них уровне.

Для создания интеллектуальной информационно-аналитической системы оценки состояния здоровья студента предполагается сбор больших объемов данных. В качестве объекта исследования были взяты студенты 1 курса Казахского национального университета имени Аль-Фараби. Перед проведением опроса HBSC среди студентов 1 курса было получено разрешение от авторов и учреждений, составлявших опросные вопросы. Опрос состоит из закрытых вопросов с предварительно закодированными ответами и является максимально коротким, чтобы увеличить процент ответов и обеспечить полное сохранение выборки. Последняя версия анкеты была переведена на казахский и русский языки, а затем снова переведена на английский. Основные разделы анкеты: социальные показатели по отношению к семье, сверстникам и университету; показатели по отношению к здоровью и настроению; формы поведения по отношению к здоровью; формы поведения, связанные с риском для здоровья. Сбор данных осуществлялся в системе «Univer».

Опрос HBSC состоит из 2 этапов:

- На 1 этапе проводится анкетирование студентов через систему «Univer». Этот опрос помогает определить психологическое, социальное, а также состояние здоровья студентов.
- На 2 этапе анкета берется анонимно, так как в системе “Univer” студенты могут не дать правдивых ответов на эти вопросы.

Как уже говорилось, рассмотрены все факторы, влияющие на состояние здоровья студента. Например, в опроснике представлен список из 8 вопросов, которые позволяют выявить жалобы студента на проблемы со здоровьем. По исследованиям специалистов, данные вопросы позволяют выявить заболевания, которые могут быть на начальном этапе подросткового возраста, а в запущенных случаях могут иметь свое продолжение и во взрослом возрасте.

Каждый показатель может быть тесно связан с другими и оказывать свое влияние. Например, на правильное или нездоровое

питание студента, массу тела, желание жить, жалобы на здоровье и другие показатели может повлиять благополучие его семьи. Это связано с тем, что из-за низкого социально-экономического уровня он может не иметь возможности потреблять необходимые ему виды продуктов. В ходе проведения аналитической работы выявляются не только тематические группы анкет, отражающие их связь с социально-экономическим показателем, но и взаимосвязь между другими аспектами.

#### ***1.1.4 Цифровые платформы по контролю и профилактике заболеваний молодежи***

В связи с тем, что сеть исследований поведения детей школьного возраста и студентов по состоянию здоровья (HBSC) реализуется на международном уровне, существуют различные платформы, которые анализируют и визуализируют собранные данные. Наиболее популярными из существующих на мировом уровне платформ для анализа и визуализации полученных данных являются американская платформа «центры по контролю и профилактике заболеваний», Европейский медицинский информационный портал.

«Центры по контролю и профилактике заболеваний» являются платформой для защиты американских граждан от рисков для здоровья, безопасности физического и психического здоровья. На платформе представлены данные Национальной, государственной и местной системы наблюдения за рискованным поведением молодежи (YRBSS), полученные от граждан США в виде таблиц, диаграмм, карт.

Система YRBSS контролирует шесть категорий здоровья и сопутствующего поведения, которые способствуют основным причинам смертности и инвалидности среди молодежи и взрослых, включая:

- поведение, способствующее неосторожной травме и насилию;
- сексуальное поведение из-за непредвиденной беременности и заболеваний, передающихся половым путем;
- употребление алкоголя и других наркотиков;

- использование табака;
- диетическое вредное поведение;
- недостаточная физическая активность.

YRBSS также изучает распространенность ожирения и астмы и другое поведение, связанное со здоровьем.

YRBSS-система опросов. Он включает в себя 1) CDC и национальный опрос, проводимый на основе местных опросов, проведенных государственными, территориальными организациями и 2) штатными, территориальными и местными органами образования и здравоохранения и племенными правительствами.

Анализ результатов системы YRBSS на платформе CDC представлен в виде таблиц, диаграмм, карт. Платформа CDC показывает сравнительный статистический анализ результатов опроса по полу, возрасту между штатами, территориями (Рис. 1.2). То есть, если какой-либо вопрос анкеты анализируется по полу, то указывается процент их ответов.

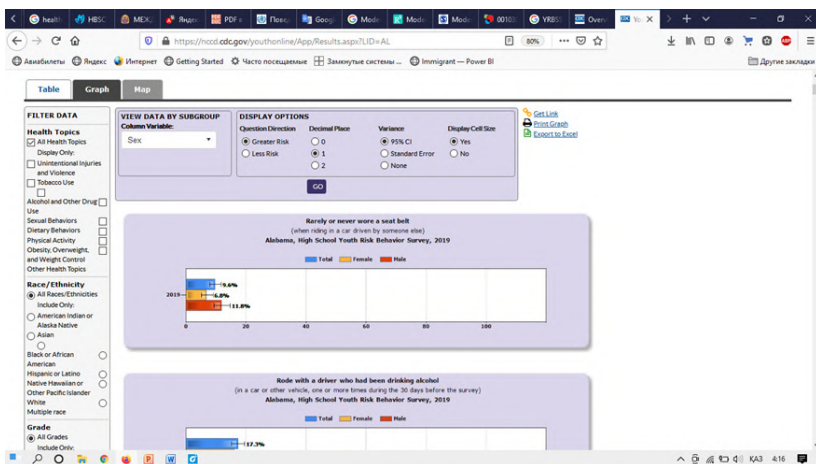


Рисунок 1.2 - Аналитическая работа, выполненная на платформе CDC

«Центры по контролю и профилактике заболеваний» обеспечены онлайн-веб-порталом для эпидемиологических исследова-

ний под названием Wonder<sup>38</sup>. CDC WONDER управляет около 20 общедоступными наборами данных о рождении, смерти, диагностике рака, туберкулезе, вакцинации, воздействии на окружающую среду и оценке населения и многих других темах в США. Эти наборы данных доступны в виде онлайн баз данных, которые предоставляют общедоступный доступ к специальным запросам, сводной статистике, картам, диаграммам и извлечению данных. Большая часть данных обновляется каждый год; некоторые коллекции обновляются ежемесячно или еженедельно. Большинство из этих ресурсов имеют удобные для использования инструменты запросов, а некоторые из них предоставляют данные на государственном уровне по всему штату<sup>39</sup>. Данные готовы к использованию в настольных приложениях, таких как текстовые процессоры, табличные программы или пакеты статистического и географического анализа. Доступные форматы файлов включают простой текст (ASCII), веб-страницы (HTML) и файлы электронных таблиц. Коллекции, доступные в CDC WONDER, основаны на сотрудничестве нескольких агентств. Многие программы CDC сотрудничают с CDC WONDER и предоставляют доступ к информации внешним партнерам. Государственные, местные и региональные департаменты здравоохранения полагаются на данные, предоставленные CDC WONDER, которые предоставляют обзор состояния здоровья сообщества, оценивают работу программы с целью планирования и сравнивают свое сообщество с другими местами.

Andrew Friede в своем исследовании «Центры по контролю и профилактике заболеваний» рассмотрел подходы, заложенные в создании онлайн-портала Wonder. Все данные хранятся и обновляются в основном кадре CDC<sup>40</sup>. Американские «Центры по контролю и профилактике заболеваний» не собирают информацию о контроле за здравоохранением напрямую, но полагаются на госу-

---

<sup>38</sup> “WONDER online databases utilize a rich ad-hoc query system for the analysis of public health data” [Электронный ресурс]. – URL: <https://wonder.cdc.gov/>

<sup>39</sup> “CDC WONDER Online Databases” [Электронный ресурс]. – URL: <https://wonder.cdc.gov/DataSets.html>

<sup>40</sup> Friede A., Daniel H. Rosen, Joseph A. R., Journal of the American Medical Informatics Association – 1994. – Vol.1(4), – P. 303–312.

дарственные и местные департаменты здравоохранения и другие системы для этого. В течение последних нескольких лет наблюдатели внутри и за пределами CDC обнаружили некоторые важные факторы, влияющие на контроль в 21 веке (такие как проблемы безопасности, технологические достижения и реформа здравоохранения), и то, как эти факторы могут повлиять на контрольное предприятие. Наблюдатели отметили необходимость постоянной оценки систем контроля; стандартизация с целью развития устойчивых и интегрированных систем; адаптация систем и рабочей силы к современным требованиям. Эти наблюдатели признали многие проблемы, которые могут помешать прогрессу, такие как проблемы с финансированием, рабочей силой, стандартами информационных технологий, конфиденциальностью пациентов и заботой о доступе к данным, качестве и обмене данными. В 2014 году «центры по контролю и профилактике заболеваний» запустили новую стратегию мониторинга, чтобы облегчить и улучшить возможности мониторинга Агентства и системы здравоохранения. Эта новая стратегия нацелена на устранение нежелательной информации в сборе данных надзора и отчетности агентства, а также на повышение функциональной совместимости систем IT-контроля<sup>41</sup>. Целью этой инициативы является повышение эффективности сбора данных, а также обеспечение доступности и совместимости данных национального мониторинга. CDC путем создания новых служб данных:

- направленных на улучшение интеграции независимых, изолированных систем;
- снижение бремени отчетности для государств;
- сокращение времени на обнаружение данных, представляющих интерес для национального контроля, или реагирование на чрезвычайные ситуации (например, расследование эпидемии);
- предоставление ценных данных о контроле над программой CDC;
- обеспечение быстрых действий в области здравоохранения.

---

<sup>41</sup> “Public Health Surveillance, Preparing for the Future” [Электронный ресурс]. – URL: <https://www.cdc.gov/surveillance/pdfs/Surveillance-Series-Bookleth.pdf>

«Центры по контролю и профилактике заболеваний» создают эти услуги, как показано на рисунке 1.3, чтобы лучше связать CDC с растущим количеством услуг и систем, которые могут предоставлять данные о проблемах со здоровьем из больниц, кабинетов врачей, аптек и других источников здравоохранения<sup>42</sup>.

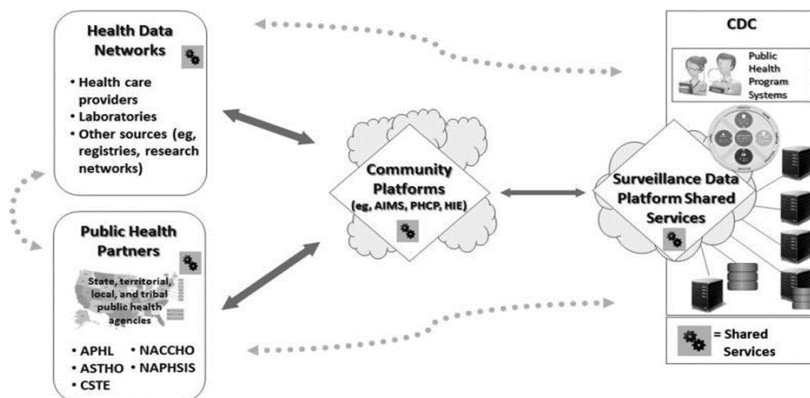


Рисунок 1.3 – Перенос данных между CDC и источниками данных

На рисунке 1.3 показаны данные, полученные от партнеров в области здравоохранения и общественного здравоохранения, передаваемые через платформы сообщества общественного здравоохранения, поступающие на платформу данных мониторинга CDC и общие услуги и, в конечном счете, передаваемые в соответствующие системы, базы данных и программы CDC. Общие услуги регулируются на основе надлежащего использования данных для нужд программ здравоохранения.

Портал Европейской ассоциации общественного здравоохранения (The European Public Health Association) является еще одним из успешных проектов мониторинга состояния здоровья населения<sup>43</sup>. ВОЗ следит за региональными и глобальными условиями и тенденциями в области здравоохранения, интегрирует все болезни и ин-

<sup>42</sup> “Public Health Surveillance and Data” [Электронный ресурс]. – URL: [https://www.cdc.gov/surveillance/projects/dmi-initiative/where\\_does\\_our\\_data\\_come\\_from.html](https://www.cdc.gov/surveillance/projects/dmi-initiative/where_does_our_data_come_from.html)

<sup>43</sup> [Электронный ресурс]. – URL: <https://eupha.org/>

формационные системы здравоохранения. Портал Европейской ассоциации общественного здравоохранения создан как универсальный журнал медицинской информации, отвечающий различным потребностям и навыкам политиков, экспертов, сотрудников ВОЗ и общества в Европейском регионе ВОЗ. Он предлагает инновационный подход к представлению ключевой информации в области здравоохранения, предложенный Европейским региональным бюро ВОЗ. Доступ к данным ВОЗ Европы осуществляется с помощью API хранилищ данных. Программа хранилища данных API используется для обеспечения программного доступа к содержимому хранилища данных в Европейском региональном офисе ВОЗ.

К рекомендуемым техническим строительным блокам портала относятся:

- Хранилище данных, его API (интерфейс прикладного программирования) и примеры инструментов, которые специалисты по данным могут использовать для быстрой визуализации на основе данных, полученных из API;

- Портал и его различные компоненты для представления и использования информации;

- Способы и примеры интерактивного инфографического дизайна и передачи истории данных;

- Разработка и использование мобильных приложений для распространения медицинской информации.

- API можно использовать для поиска недавно добавленных наборов данных, для поиска в хранилище данных, для анализа содержимого хранилища данных, для загрузки или для поиска индикаторов и наборов данных, связанных с общей темой. С помощью этого API можно запросить более 8000 показателей, используя запросы нескольких типов с различными параметрами. Как показано на рисунке 1.4, каждый источник поступает в хранилище данных из различных источников данных, в том числе из информационных систем, социальных организаций. Эти данные анализируются и передаются через портал Европейской ассоциации общественного здравоохранения различными инструментами в виде диаграмм, таблиц, карт и т.д.<sup>44</sup>

<sup>44</sup> “Digital technologies for key public health functions: results of an ECDC expert consultation”

На портале наглядно представлены результаты опроса HBSC, проведенного в странах-членах ВОЗ. API хранилища данных позволяет находить вновь добавленные наборы данных, искать данные, загружать данные для анализа или искать размеры и наборы данных, объединенные общей темой.

Общая структура модели данных:

1. показатель (критерий) – это ряд данных, которые служат заполнением того или иного показателя или другой измеряемой категории. Другими словами, индикатор содержит набор значений этого показателя для всех стран, для которых данный показатель определен, или набор ответов этих стран на отдельный вопрос опроса. HBSC показывает содержание каждого вопроса анкеты, полное название.

2. каждый показатель сопровождается пояснениями, в которых дается его определение, а также оговорками по странам, в которых приводится информация о специфике данных по той или иной стране.



Рисунок 1.4 - Средства доступа к данным ВОЗ Европы



3. ряд показателей можно объединить в набор данных. Каждый набор данных содержит метаданные и свойства (владелец данных и т. д.).

4. показатели по смежным темам могут быть объединены в тематические классификации. Классификации относятся к различным наборам данных, но позволяют запрашивать тематически связанные показатели. На рисунке 1.5 представлена аналитическая работа портала Европейской ассоциации общественного здравоохранения по результатам опроса HBSC из стран Европы.

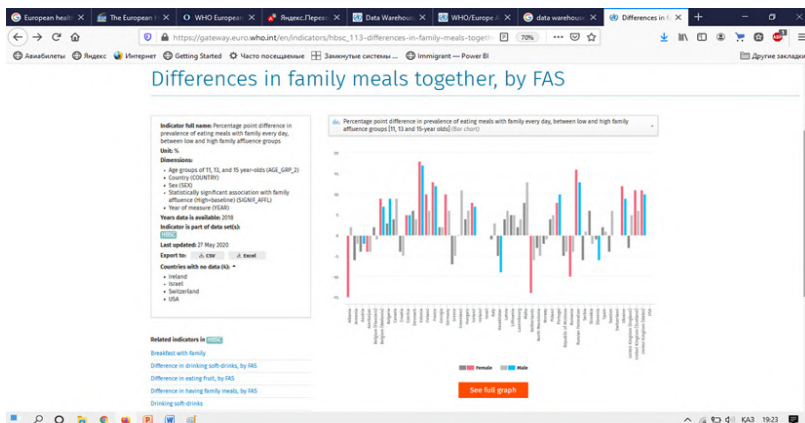


Рисунок 1.5 - Визуализация на европейском информационном портале здравоохранения

Анализ результатов опроса HBSC представлен в виде визуализации. На рисунке полностью представлена общая структура модели данных, то есть сведения о показателе расположены слева от диаграммы, а комментарии и примечания к ним - внизу диаграммы. Кроме того, экспортируемые данные и графики выполняются с помощью файла Excel. Диаграмма на данном рисунке описывает специфику показателя завтрака с семьей по социально-экономическому параметру. Здесь проводится анализ по возрасту и полу респондентов.

Таким образом, можно сделать вывод, что разработка информационно-аналитической системы на основе опросников с инструментами анализа и визуализации данных позволит получить объективную картину состояния здоровья студентов университета с точки зрения статистики.

Одним из больших преимуществ студенческого контингента КазНУ им. Аль-Фараби является то, что в контингент входят молодые люди со всех регионов Казахстана<sup>45</sup> (Рис. 1.6), что позволяет рассматривать КазНУ как образец для обучения и управление здоровьем студенческой молодежи по всей стране.



Рисунок 1.6 – Представленность различных регионов РК в контингенте студентов КазНУ

Целью данного исследования является оказание помощи в разработке и реализации будущих стратегий и мер в области цифрового здравоохранения в университетских городках. Это предварительное исследование было направлено на предоставление обзора моделей использования цифрового здравоохранения среди студентов университетов в Казахстане, расширение существующих исследований обновленными данными. Также изуча-

<sup>45</sup> [Электронный ресурс]. – URL: <https://eupha.org/> <https://www.kaznu.kz>

ется взаимосвязь антропоморфологических данных с социально-демографическими характеристиками и самооценкой здоровья.

## **1.2. Техническое описание интеллектуальной информационно-аналитической системы**

Архитектура разрабатываемой интеллектуальной информационно-аналитической системы состоит из следующих основных компонентов:

1. модуль сбора, хранения и предобработки (структурирование и первичная очистка) данных по показателям здоровья студентов;
2. хранилище данных с общей информацией о студентах и состоянии их здоровья;
3. модуль предобработки данных для гибкого использования методами интеллектуального анализа;
4. модуль интеллектуального анализа данных комплекса показателей здоровья студентов;
5. модуль визуализации результатов анализа данных;
6. модуль поддержки принятия решений и выработки рекомендаций.

### ***1.2.1 Источники данных***

Проведение анализа оценки состояния здоровья производится путем создания хранилища данных, при этом источниками информации являются:

- результаты исследований физического развития и состояния здоровья студентов, полученных в результате проведения онлайн-опросов;
- результаты исследований физического развития и состояния здоровья студентов, полученных в результате медицинского осмотра;
- данные о студентах из системы управления учебным процессом (система «Univer»);
- данные показателей загрязнения воздуха, состояния окружающей среды и др.

Сбор первичной информации проведен в виде одномоментного поперечного исследования показателей здоровья и показателей, связанных со здоровьем. Объектом исследования определены студенты 1 курса всех специальностей КазНУ возрастного диапазона 17-19 лет. Данный возраст еще входит подростковую группу согласно классификации ВОЗ, и находится на границе перехода во взрослую возрастную группу (в Казахстане молодые люди с 18 лет относятся к взрослой возрастной группе). ВОЗ рассматривает подростковый возраст как «уникальный этап человеческого развития и важное время для создания основ хорошего здоровья»<sup>46</sup>. На этом этапе подростки устанавливают модели поведения, например, связанные с питанием, физической активностью, употреблением психоактивных веществ и сексуальной активностью, которые в значительной степени будут влиять на их здоровье в будущем.

Исследуемая группа характеризуется также тем, что молодые люди находятся в ситуации кардинального изменения окружающей среды (социальной, физической, физиологической и т.д.), связанной со вступлением в студенческую жизнь. Данная группа рассматривается как исходная точка для дальнейших исследований психофизиологических адаптаций и развития, связанных с упомянутыми изменениями окружающей среды.

В рамках исследований оценивается текущее состояние здоровья исследуемого контингента с последующим множественным регрессионным анализом показателей в контексте факторов окружающей среды в предшествующий период жизни.

Основные методы сбора первичных данных: опрос и медицинский осмотр с проведением физикальных и неинвазивных инструментальных методов оценки физиологического статуса. Был использован стандартизованный опросник ВОЗ для оценки показателей, определяющих основные предотвратимые причины смертности и инвалидизации исследуемой возрастной группы, таких как восприятие собственного тела, буллинг и драки, пищевое поведение, жалобы на здоровье, травмы, удовлетворение

---

<sup>46</sup> [Электронный ресурс]. – URL: [https://www.who.int/health-topics/adolescent-health#tab=tab\\_1](https://www.who.int/health-topics/adolescent-health#tab=tab_1)

жизнью, ожирение, здоровье ротовой полости, физическая активность и малоподвижный образ жизни, отношения в семье и со сверстниками, школьная среда, самооценка здоровья, сексуальное поведение, социально-экономическая среда, употребление психоактивных веществ: алкоголь, табак и каннабис, снижение веса<sup>47</sup>. Методология опроса, интерпретация результатов проводится согласно руководству ВОЗ по данному инструменту опроса. С целью выявления детерминант здоровья, использована унифицированная классическая методика изучения и оценки физического развития детей и подростков, включающая измерение роста стоя (длины тела – ДТ), веса (массы тела – МТ) и окружности грудной клетки (ОГК), в паузе, на вдохе и выдохе, а также окружности талии и бедер (ОТ, ОБ). В группе физиометрических показателей физического развития исследованы жизненная ёмкость лёгких (ЖЕЛ), показатель силы сжатия кисти, становая сила, систолическое (САД) и диастолическое (ДАД) артериальное давление, частота сердечных сокращений (ЧСС). А также дана оценка неврологического и офтальмологического статуса.

Для первичной регистрации, описания и классификации характеристик здоровья и факторов окружающей среды использована Международная классификация функционирования, ограничений жизнедеятельности и здоровья детей и подростков (МКФ-ДП)<sup>48</sup> и Международная классификация функционирования, ограничений жизнедеятельности и здоровья (МКФ)<sup>49</sup>, что позволяет обеспечить глобальную концептуальную рамку и международный стандартизированный терминологический язык при классификации полученных данных. МКФ включает все аспекты здоровья человека и некоторые составляющие благополучия, относящиеся к здоровью, включая факторы окружа-

<sup>47</sup> Robert C., Freeman J., Samda O., Schnohr C. W., de Looze M. E., Nic Gabhainn S., Iannotti R., Rasmussen M., International HBSC Study Group (2009), The Health Behaviour in School-aged Children (HBSC), study: methodological developments and current tensions // *International journal of public health*, – Vol.54(2), – P.140–150. doi:10.1007/s00038-009-5405-9

<sup>48</sup> [Электронный ресурс]. – URL: <https://www.who.int/publications/list/2016/icd-children/ru/>

<sup>49</sup> [Электронный ресурс]. – URL: <https://apps.who.int/iris/bitstream/handle/10665/85389/924454542X.pdf?sequence=1&isAllowed=y>

ющей среды, описывая их в терминах доменов здоровья и доменов, связанных со здоровьем. Кодирование проводится с использованием универсального определителя по МКФ, соответствующего степени нарушений функций и структуры организма, ограничений активности в диапазоне от 0 (нет нарушений, незначительные нарушения) до 4 (полное, абсолютное нарушение), наличие препятствия или барьера со стороны окружающей среды (0 - нет барьеров до 4 – абсолютное препятствие/барьер), или присутствие облегчающих факторов (0 - нет облегчения до +4 полная компенсация).

Протокол исследования одобрен Локальной этической комиссией факультета медицины и здравоохранения КазНУ имени аль-Фараби. Участие студентов в исследовании добровольное и проводится на основе полного информированного согласия на участие в исследовании в соответствии с Хельсинкской декларацией Всемирной медицинской ассоциации. Перед проведением опроса и медицинского осмотра получено информированное согласие участника (или его законного представителя) на сбор, хранение, обработку, предоставление доступа к персональным медицинским данным отдельным оговоренным категориям работников университета, интеграцию данных в систему “Универ”, возможную передачу персональных медицинских данных в национальную систему здравоохранения. Обеспечена возможность и информированность участников (или их законных представителей) о возможности предоставлении доступа к персональным данным по запросу, отслеживании журнала доступа и отзыва согласия на любом этапе исследования и дальнейшего функционирования информационной системы с выбором опций полного уничтожения или использования только в обезличенной форме в исследовательских целях. Гарантируется конфиденциальность всех собираемых и архивируемых данных. Участникам присваиваются идентификационные номера, и во всех реестрах и электронных базах указываются только эти номера. Предусмотрен протокол шифрования, обеспечения сохранности, предоставления декларированного доступа и уничтожения персональных медицинских данных.

Вопросы соблюдения конфиденциальности очень важны при разработке систем анализа состояния здоровья<sup>50</sup>. Владельцы аккаунтов (студенты) могут видеть свои персональные медицинские данные и результаты исследования. Доступ к персональной медицинской информации для медицинских работников университета, включая членов исследовательской группы, осуществляется на условиях положений Кодекса Республики Казахстан «О здоровье народа и системе здравоохранения» от 7 июля 2020 года № 360-VI ЗРК. Немедицинский персонал университета и члены исследовательской группы имеют доступ только к обезличенным данным. При необходимости агрегации и извлечения персональных медицинских данных при оговоренных условиях и для оговоренных категорий лиц (немедицинских работников) происходит сопоставление идентификационных номеров с персоналиями студентов на основе запроса доступа к владельцу аккаунта (студента).

Таким образом, разрабатываемая интеллектуальная информационно-аналитическая система оценки состояния здоровья студентов позволит получать информацию о состоянии здоровья студентов, определять группы студентов, которым требуется помощь специалистов, оценивать воздействие проведения профилактических мероприятий по поддержке здоровья студентов, прогнозировать состояние здоровья студентов.

### ***1.2.2 Организация хранилищ данных***

Для организации хранилища данных применен трехфазный процесс ETL («Extraction – Transformation – Load» – «Извлечение – преобразование – загрузка»), который позволяет синтаксически и семантически привести исходные данные в соответствие со структурой и терминологией целевого хранилища данных. При этом в хранилище реализованы функции CDA (Complete Data Automerge). Так как данные могут быть извлечены из баз данных, управляемых различными приложениями, развернутыми на сто-

---

<sup>50</sup> Hoens T.R., Blanton M., Steele A., Chawla N.V., Reliable medical recommendation systems with patient privacy // *ACM Trans. Intell. Syst. Technol.* – 2013. – Vol. 4(4). – P. 1–31.

роне вуза, студенческой поликлиники и различных государственных статистических органов, загрузка таблиц в CDA-хранилище включает два действия.

*Создание таблиц для промежуточных данных* – действие, при котором данные считываются из базы данных конкретного приложения и записываются в общую промежуточную таблицу хранилища. Это действие называется извлечением, зависящим от источника (SDE – source-dependent extract). Записи в промежуточной таблице имеют общий формат, независимо от источника, из которого они были извлечены.

*Загрузка, не зависящая от источника* – действие, при котором данные считываются из промежуточной таблицы и записываются в конечную таблицу хранилища. Это действие называется (SIL – Source-Independent Load) – загрузкой, не зависящей от источника.

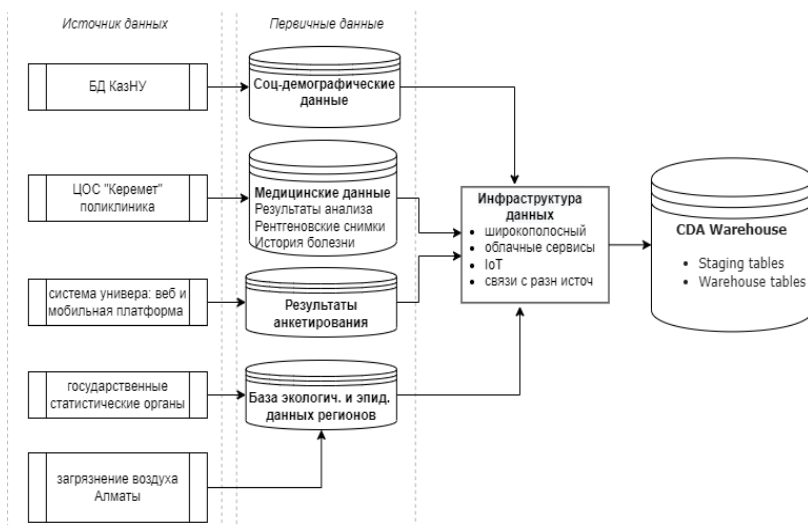


Рисунок 1.7 - Архитектура ETL

Процесс ETL (Рис. 1.7) для поддержки синхронизации данных обычно состоит из двух последовательных шагов. На этом этапе эксперты в предметной области по исходным данным определяют соответствующие элементы данных для извлечения, не-



обходимые для заполнения целевой базы данных, и определяют сопоставления между исходными данными и целевыми элементами данных. В разрабатываемой системе ETL выполняется путем отображения информации, и этот процесс организован с использованием простых сценариев Python, которые размещаются внутри отдельных микросервисов.

Преобразователь (Transformation) – это программа, которая выбирает данные из одной или нескольких таблиц, выполняет преобразования и получение данных, а также вставляет или обновляет результаты в целевой таблице. Система использует планирование задач для обновления и записи данных путем запуска всего набора ETL, необходимого для заполнения CDA-хранилища.

В системе используются 3 типа данных:

- реляционные;
- файлы;
- временные ряды.

Для сбора и хранения данных использована технология MongoDB Atlas Data Lake. Технология MongoDB Atlas Data Lake предоставляет сервисы бессерверного масштабируемого озера данных. С помощью данной технологии можно комбинировать и анализировать текущие и исторические данные без перемещения данных или операционных издержек, а расходы только за выполнение запросов. Разрабатываемая система оперирует данными, полученными из разных гетерогенных источников, начиная от неструктурированных данных (текстовых) до временных рядов (time series). Таким образом, система поддерживает две технологии хранения данных: MongoDB и PostgreSQL. MongoDB – это база данных документов, а PostgreSQL – это система управления реляционными базами данных. В системе используются миграция в объектно-реляционные системы управления базами данных и конверсия реляционной базы данных в объектно-ориентированную.

На рисунке 1.8 представлена диаграмма потока данных (DFD) на уровне контекста информационно-аналитической системы.

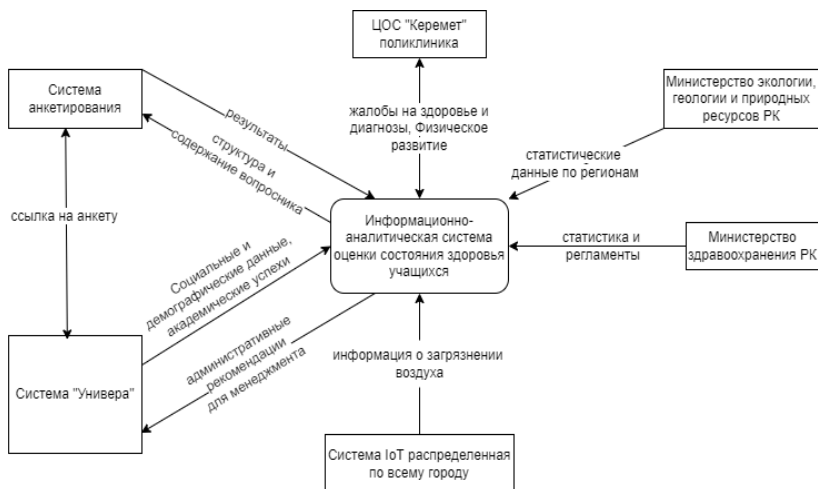


Рисунок 1.8 - Диаграмма потока данных на уровне контекста системы

### 1.2.3 Модуль анкетирования студентов

Для исследования вопросов физического развития и здоровья студентов разработаны опросники по показателям здоровья и факторам риска развития заболеваний на основе международной классификации функционирования, ограничений жизнедеятельности и здоровья.

К положительным аспектам характеристики здоровья относятся факторы, облегчающие функциональную и структурную целостность, активность и окружающая среда, к отрицательным относятся нарушения функционирования организма, ограничение активности и возможности участия, факторы и барьеры внешней среды. Функционирование человека в определенной области представляется как взаимодействие или сложное соотношение между изменениями здоровья и контекстными факторами (факторами окружающей среды и индивидуальными факторами). Между этими элементами существует динамическое взаимодействие: вмешательства на уровне одного элемента могут изменять другой или другие элементы.

Основной задачей исследования является международная классификация функционирования классификация с помощью методов искусственного интеллекта.

Для проведения медицинского осмотра и организации опросов разработан стандартизированный протокол исследования, обеспечивающий теоретическую основу тем исследований и процедур сбора и анализа данных. Протокол направлен на обеспечение безопасности сопоставимых данных. С целью сбора больших объемов данных, необходимых для создания интеллектуальной информационно-аналитической системы оценки состояния здоровья студента через платформу «Univer» проведен опрос студентов 1 курса Казахского национального университета имени Аль-Фараби (Рис. 1.9), в котором приняло участие около 900 студентов.

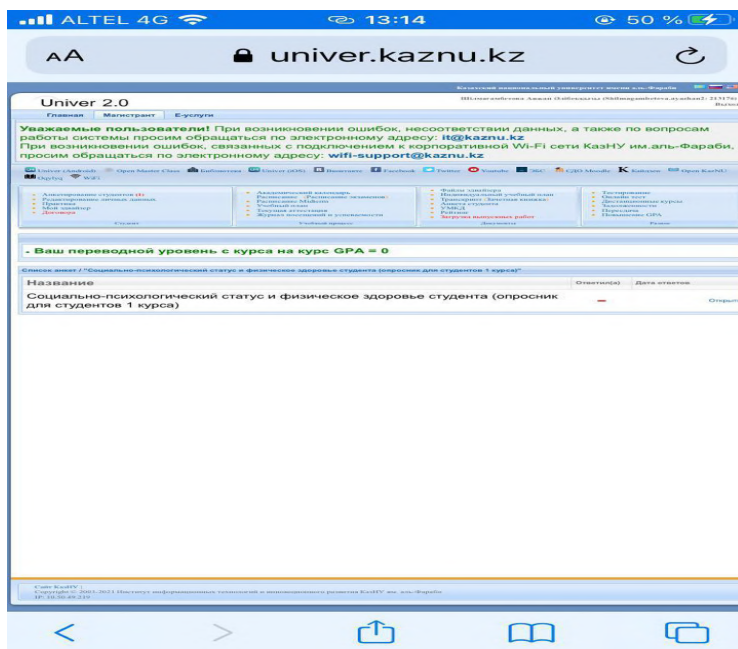


Рисунок 1.9 - Страница с опросником в системе Универ

Перед проведением опроса HBSC среди студентов 1 курса был получен разрешение от авторов и учреждений, составлявших опросные вопросы. Опрос состоит из закрытых вопросов с предварительно закодированными ответами и является максимально коротким, чтобы увеличить процент ответов и обеспечить полное сохранение выборки. Последняя версия анкеты была переведена на казахский и русский языки. Основные разделы анкеты: социальные показатели по отношению к семье, сверстникам и университету; показатели по отношению к здоровью и настроению; формы поведения по отношению к здоровью; формы поведения, связанные с риском для здоровья. Сбор данных осуществлялся в системе «Univer».

Опрос HBSC состоит из 2 этапов:

– проводится опрос студентов через систему «Univer». Этот опрос помогает определить социальное положение, а также физическое и психологическое состояние здоровья студентов.

– опрос берется анонимно, так как в системе «Univer» студенты могут не дать правдивых ответов на эти вопросы.

Данный модуль включает следующие функции:

1. Разработка модуля для проведения онлайн опросов.
2. Сохранение данных с регулярно проводимых опросов в облачном хранилище данных.

#### ***1.2.4 Модуль интеллектуального анализа данных***

Для интеллектуального анализа данных используются следующие методы и технологии:

– Power BI - инструмент бизнес-аналитики для настольных и облачных приложений;

– методы машинного обучения.

В набор Power BI включены инструменты подготовки данных, которые значительно упрощают обычно трудоемкий процесс моделирования данных (Рис. 1.10-1.11). Power BI предлагает некоторые сложные инструменты искусственного интеллекта, включая прогнозирование и кластеризацию наборов данных.

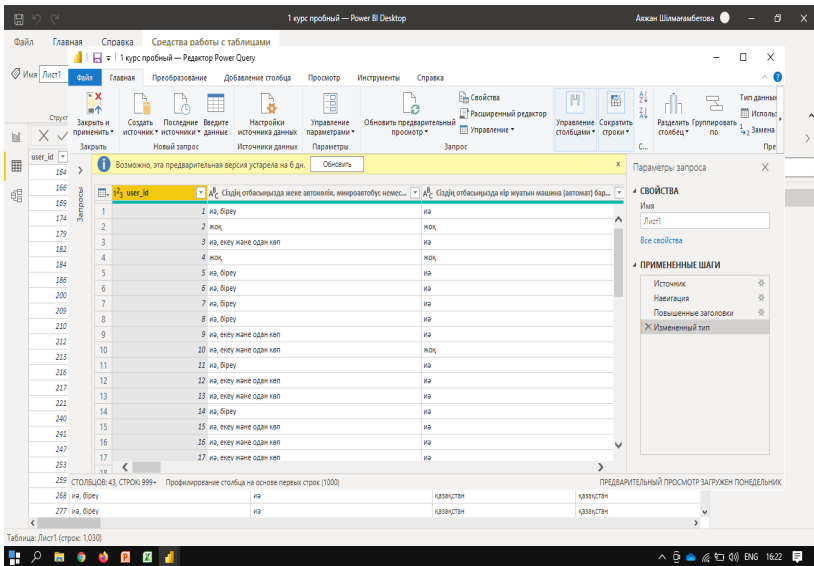


Рисунок 1.10 - Анализ данных с помощью инструмента Power BI

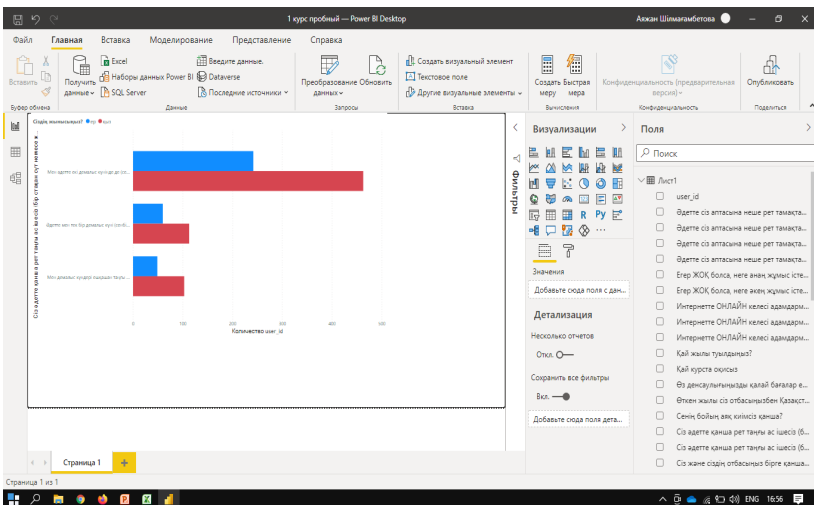


Рисунок 1.11 - Пример результатов анализа

Проведена апробация разработанных опросников и пилотного цифрового модуля с участием студентов 1-го курса (набора 2020 г.), при этом:

- Опросник HBSC (ВОЗ) оцифрован, разделен на анкеты №1 и №2.
- Опросник №1 интегрирован с профилем студента в системе Univer и заполняется онлайн.
- Опросник №2 заполняется при медицинском осмотре.
- Медицинское обследование проводят сотрудники Высшей школы медицины КазНУ им. аль-Фараби.

### 1.2.5 Модуль визуализации данных

Данные, находящиеся в хранилище, используются в качестве входных данных для дальнейшего анализа, например, для классификации, кластеризации с применением методов машинного обучения или другой обработки данных. Кроме того, система предоставляет инструменты статистического, корреляционного анализа и визуализации на основе технологии Power BI.

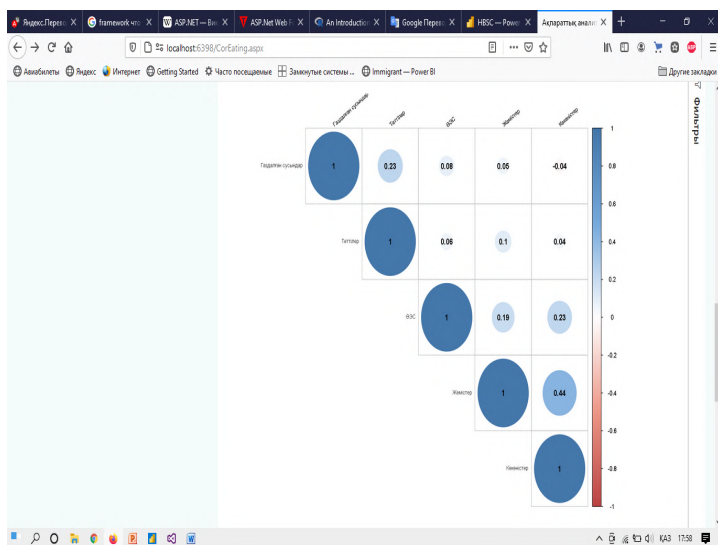


Рисунок 1.12 - Корреляционная связь между типами НЭС и питания

Всего в опросе HBSC приняли участие 890 первокурсников, из них 590 девочек и 300 мальчиков 2001-2006 годов рождения. Можно обсудить результаты некоторых аналитических работ. Корреляционная связь между социально-экономическим уровнем и видами питания представлена на рисунке 1.12.

Как видно из рисунка, существует положительная слабая корреляция между социально-экономическим уровнем студента и потреблением здоровой пищи, то есть фруктов и овощей, в отличие от употребления вредной пищи. Среди вредных продуктов можно наблюдать слабую корреляцию, а среди полезных – среднюю. А между употреблением здоровой и вредной пищи можно увидеть негативную связь, делается вывод, что они не имеют к друг другу никакого отношения.

На рисунке 1.13 показан уровень здорового питания по параметру ВЭС.

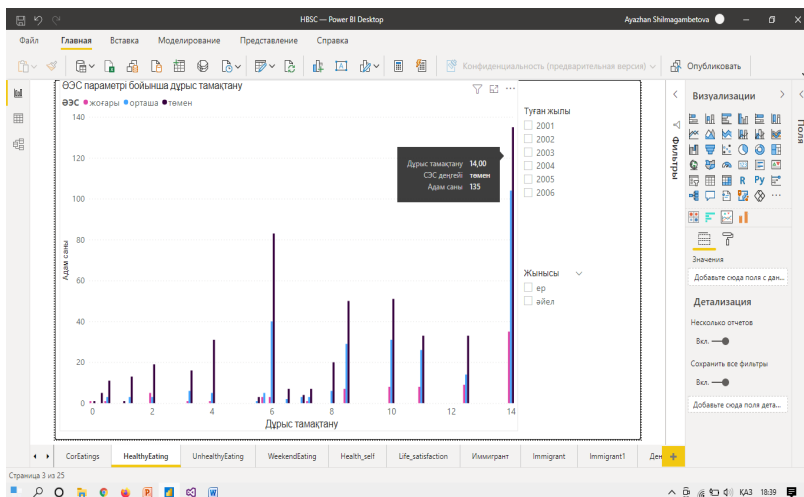


Рисунок 1.13 - Правильное питание по параметру ВЭС

Потребление фруктов и овощей было объединено в категорию здорового питания. Эти два пункта были объединены в индекс в диапазоне от 0 до 14, 0 соответствует тому, что они никогда не потребляют фрукты и овощи, а 14 соответствует

тому, что они потребляют фрукты и овощи по крайней мере один раз в день. Результаты показывают связь между социально-экономическим уровнем семьи и ежедневным потреблением вредной пищи. Как показано на рисунке 1.13, пользователь может получить необходимую ему информацию. Значение пола и возраста принято принимать как независимую переменную, а социально-экономический показатель здорового питания и школьника – как зависимую переменную. Пользователь может получить информацию, установив для себя желаемый возраст и пол. При расчете ВЭС было установлено, что из 890 студентов 522 студентов имели низкую ВЭС, из 283 студентов – среднюю, а из 85 студентов – высокую ВЭС. В целом, если сделать вывод из таблицы, то можно заметить, что уровень здорового питания очень высок, даже при низком уровне ВЭС, подобное явление наблюдается у 135 студентов, в том числе у 49 мальчиков и у 86 девочек, что является самым высоким среди других уровней здорового питания. Аналогичное явление наблюдается и у студентов со средним и высоким социальным уровнем, т. При высоком уровне потребления здоровой пищи количество людей со средним и высоким СОЭ больше, чем при других уровнях здорового питания. 14 уровню питания соответствуют в среднем 104 студента ВЭС, выше – 35 студентов ВЭС. Можно сделать вывод, что большая часть студентов питается здоровой пищей.

А обратную картину можно наблюдать на уровне вредного питания (Рис. 1.14).

Потребление газированных напитков и сладостей было объединено в категорию вредного питания. Эти два пункта были объединены в индекс в диапазоне от 0 до 14, 0 соответствует тому, что они никогда не потребляют газированные напитки и сладости, а 14 соответствует тому, что они потребляют газированные напитки и сладости, по крайней мере, один раз в день. Результаты показывают связь между социально-экономическим уровнем семьи и ежедневным потреблением вредной пищи.



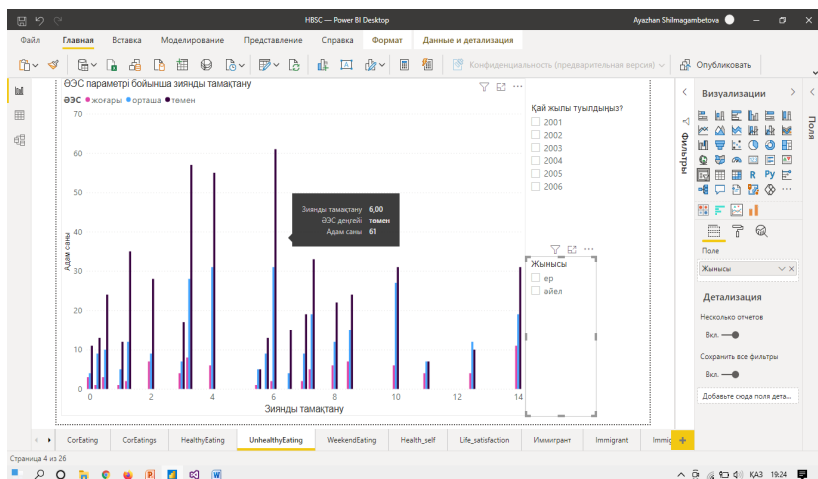


Рисунок 1.14 - Вредное питание по параметру ВЭС

Из рисунка 1.14 видно, что уровень вредного питания не сильно вырос. Наибольший показатель наблюдается между 3,5-6 вредными пищевыми продуктами. Уровень потребления вредной пищи у студентов низкий и средний. Однако можно заметить, что уровень потребления вредной пищи выше у студентов с высоким уровнем ВЭС (11 студентов, 2001-2003 годы). Благодаря высокому содержанию ВЭС можно свободно употреблять сладкие и газированные напитки.

На рисунке 1.15 показана корреляционная связь между жалобами на здоровье.

На рисунке 1.15 имеется средняя корреляционная связь между нервным ощущением, раздражительностью и плохим настроением, плохим самочувствием. То есть, если эти жалобы на одно здоровье будут у студента, то могут быть жалобы на два других здоровья. Кроме того, как показано на рисунке, средняя корреляция 0,34 и 0,31 между жалобами на головную боль и чувство головокружения и плохое самочувствие показывает, что существует сила. Если уровень жалоб на здоровье повышается, лучше обратиться к врачу.

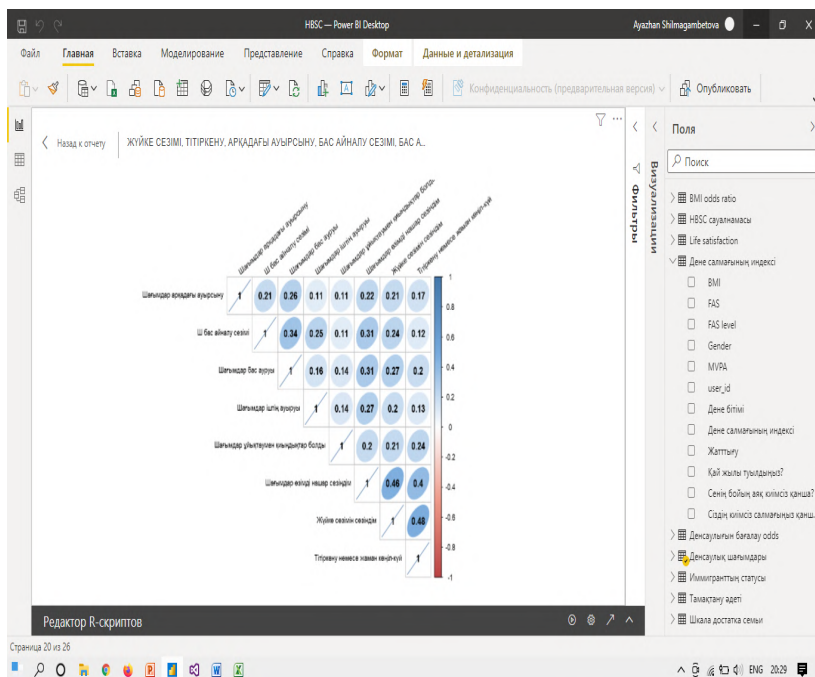


Рисунок 1.15 - Взаимосвязь между жалобами на здоровье

Масса тела – один из очень важных показателей. У него есть связь с физическими упражнениями. Некоторые люди, придерживаясь здорового образа жизни, а некоторые занимаются физическими упражнениями с целью похудения. Студентам был задан вопрос о том, сколько весят без одежды и их рост без обуви. И вне университетского времени была запрошена высокая интенсивная физическая активность, то есть время, когда студент потеет и становится активным, так что учащается сердцебиение. На рисунке 1.16 можно увидеть связь между массой тела и высокоинтенсивной физической активностью.

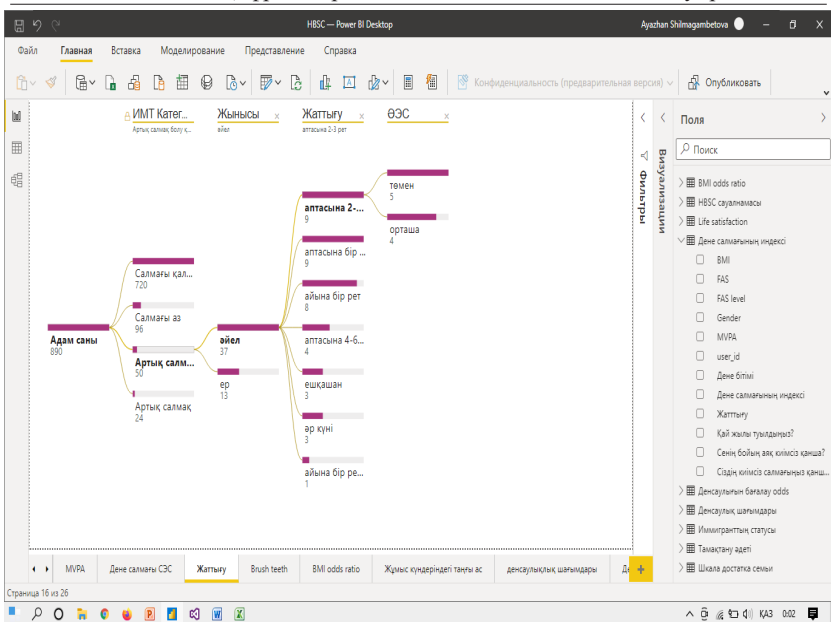


Рисунок 1.16 - Зависимость массы тела от физической активности, ВЭС

Факторный анализ можно провести с помощью панели визуализации «дерево декомпозиции», предоставленной рабочим столом Power BI. Вы можете отслеживать и определять, как каждая переменная влияет на одну независимую переменную. Визуализация позволяет разделить группу на категории, по выбранному критерию, например, факторы, влияющие на массу тела, показанную на рисунке 1.16, пользователь может использовать так, как ему это нужно. Можно заметить, что существует 37 девочек со средним и низким риском избыточного веса, которые тренируются не реже одного раза в неделю (24 студента). Это можно рассматривать как их действия с целью похудения. А лишний вес встречается у 12 девочек и 12 мальчиков. Установлено, что большая часть мальчиков занимается спортом не менее 1 раза в семь. И было установлено, что 24 студента с избыточным весом и категорией избыточного веса не являются физически активными, они не занимаются спортом один раз в месяц или никогда. Риск

избыточного веса может быть вызван отсутствием движения. Возможно, из-за карантинных мер, онлайн занятий. После того, как Университет перешел на оффлайн обучение, следует уделить большое внимание занятиям физической культурой. Это связано с тем, что избыточный вес в подростковом возрасте может привести к возникновению других заболеваний в будущем при переходе во взрослую жизнь.

При исследовании медицинских данных используется соотношение возможностей. Он часто используется в поиске связи между заболеваниями и их факторами. При проведении опроса HBSC была рассчитана связь массы тела с полом, которая была определена. На рисунке 1.17 показано допустимое соотношение массы тела.

The screenshot shows the Power BI Desktop interface with a table of BMI data and a table of Odds ratios. The table of Odds ratios is as follows:

BMI	Odds ratio	Lower 95% CI	Upper 95% CI
Артық салмақ болу қаупі	0.36	0.125	0.999
Салмағы қалыпты	0.46	0.197	1.048
Салмағы аз	1.09	0.437	2.705
Артық салмақ	1.00	NA	NA

Рисунок 1.17 - Отношение массы тела к полу

Как показано на рисунке 1.17, у 890 студентов 720 Вес тела нормальный. Риск избыточного веса и количество студентов с избыточным весом 74. одно наблюдаемое явление, у девочек больше риск избыточного веса. Причина в том, что у девочек в этом подростковом возрасте он быстро наполняется ростом различных гормонов, а у мальчиков наоборот, он быстро работает, а их тела

быстро выходят, становятся жесткими и набирают вес. Это также четко прослеживается в таблице. Из 300 мальчиков 50 имеют меньший вес. Если посмотреть на вторую таблицу на рисунке 1.17, то в категории с меньшим весом соотношение 1,09 означает, что явление с меньшим весом у мальчиков в 1,09 раза больше, чем у девочек. Но наблюдаемая зависимость, обусловленная наличием доверительного интервала между 0,437 и 2,705, статистически не значима, так как доверительный интервал входит в 1, нижняя граница доверительного интервала меньше 1.

Кроме того, студентам предлагается ответить на вопрос: «как вы оцениваете свое здоровье?». Ответы на вопросы: очень хорошо, хорошо, удовлетворительно и плохо. Первые два пункта были объединены в категорию «хорошо», а последние два пункта - в категорию «плохо». На рисунке 1.18 показано отношение высокой оценки здоровья к ВЭС и полу.

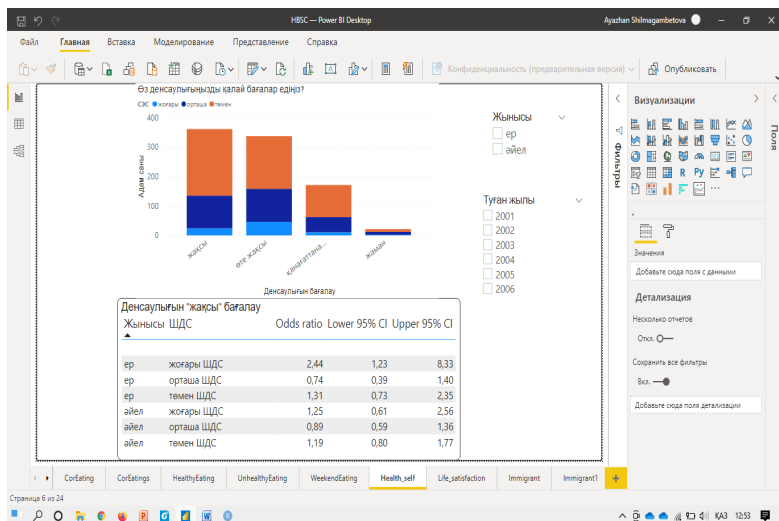


Рисунок 1.18 - Отношение оценки здоровья к ВЭС и полу

Из рисунка 1.18 видно, что большая часть студентов, в том числе 698 студентов, оценили свое здоровье как хорошее и очень

хорошее. На диаграмме видно, что студент удовлетворен своим здоровьем, даже если его семейное благополучие низкое. Если посмотреть на таблицу соотношения возможностей, то видно, что у мальчиков с высоким социально – экономическим показателем в 2,44 раза выше, чем у других невысоких мальчиков. Поскольку доверительный интервал находится между 1,23 и 8,33, наблюдаемая зависимость является статистически значимой, поскольку как нижняя граница доверительного интервала, так и верхняя граница больше 1. Когда мы обсуждаем данные в остальных строках таблицы Ад, поскольку они содержат доверительный интервал 1, наблюдаемая зависимость не является статистически значимой, поскольку нижняя граница доверительного интервала меньше 1. На рисунке 1.19 показана визуализация анализируемого соотношения возможностей по ВЭС и полу студентов, хорошо оценивших свое здоровье.

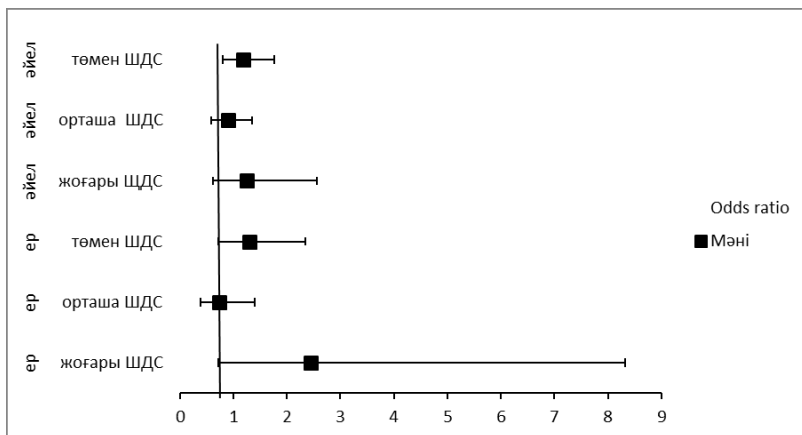


Рисунок 1.19 - Соотношение возможностей оценки «хорошо» по полу и ВЭС

Результаты анкетирования, проанализированные выше, включены в интеллектуальную информационно-аналитическую систему. Кроме того, как и выше проанализированные работы, на платформе представлены результаты всех анализируемых работ.

Также Power BI предусматривает использование AutoML модуля. С помощью Cognitive Services в Power BI можно реализовать различные алгоритмы Azure Cognitive Services для обогащения данных. Автоматическое машинное обучение (AutoML) для потоков данных позволяет бизнес-аналитикам обучать, проверять и вызывать модели машинного обучения непосредственно в Power BI.

AutoML поддерживает создание бинарных моделей прогнозирования, классификации и регрессии для потоков данных. Это типы контролируемых методов машинного обучения, которые извлекают уроки из известных результатов прошлых наблюдений, чтобы предсказать результаты следующих наблюдений. Проводить анализ данных также возможно с использованием фреймворков языка Python. Это позволяет оценить, какие данные релевантны для работы и какие параметры более значимы. Ниже приведена часть эксперимента в качестве примера. AutoML поддерживает создание двоичных моделей прогнозирования, классификации и регрессии для потоков данных. На рисунке 1.20 представлены различные инструменты анализа данных, предусмотренные в информационно-аналитической системе.

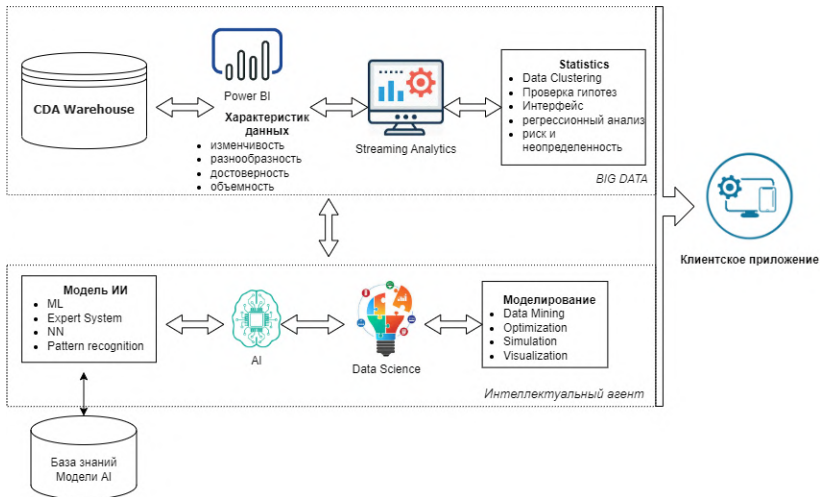


Рисунок 1.20 - Моделирование и представление данных

### ***1.2.6 Модуль поддержки принятия решений и выработки рекомендаций***

Рекомендации, основанные на цифровом профиле здоровья отдельного студента, представляют интерес для трех категорий целевых пользователей:

- самого студента;
- уполномоченных сотрудников медицинских служб университета;
- сотрудников департамента по социальным вопросам. Для реализации этого модуля будут применены алгоритмы машинного обучения.

Рассмотрены факторы, влияющие на состояние здоровья студента. В опроснике содержатся вопросы, которые позволяют выявить жалобы студента на здоровье. По исследованиям специалистов, данные вопросы позволяют выявить заболевания, которые могут быть на начальном этапе подросткового возраста, а в запущенных случаях могут иметь свое продолжение и во взрослом возрасте. Каждый показатель может быть тесно связан между собой и оказывать свое влияние. Например, на правильное или нездоровое питание студента, массу тела, желание жить, жалобы на здоровье и другие показатели может повлиять благополучие его семьи. Это связано с тем, что из-за низкого социально-экономического уровня он может не иметь возможности потреблять необходимые ему виды пищи. В ходе проведения аналитической работы выявляются не только тематические группы опросников, отражающие их связь с социально-экономическим показателем, но и взаимосвязь между другими аспектами.

По результатам анкетирования среди студентов по выявлению риска развития сахарного диабета была выявлена группа риска. В ходе исследований разработан рекомендательный функционал информационно-аналитической системы для программы управления заболеванием на основе выявления корреляционных связей между показателями уровня глюкозы в крови пациентов, физической активностью и режимом питания. Разработка функционала осуществлялась в соответствии с международными



ми рекомендациями<sup>51</sup> и стандартами оказания медицинской помощи пациентам с сахарным диабетом в РК<sup>52</sup>. Персонализация управления диетой и физической активностью позволит снизить количество эпизодов гипогликемии и добиться хорошего гликемического контроля. В персонализированную программу реабилитации пациентов с сахарным диабетом включены физическая активность и физические упражнения. При этом физическая активность (далее ФА) определяется как любое движение, увеличивающее потребление энергии, в то время как упражнения – это более систематизированный и структурированный вид физической активности. Понимание важности физической активности и физических упражнений способствует формированию у детей здоровых привычек, которые будут поддерживать их в течение всей жизни.

Ниже представлены результаты экспериментов по обучению и тестированию моделей машинного обучения. Для первоначального эксперимента мы отобрали 600 студентов, и для каждого студента был установлен специалистами из медицинской клиники показатель здоровья (*health\_level* – от 1 до 5). Список параметров выглядит следующим образом: пол, возраст, адрес, размер семьи, статус совместного проживания родителей, опекун студента, недельное учебное время, количество прошлых неудач по предмету, дополнительная образовательная поддержка, активист, хочет ли он получить высшее образование, доступ в Интернет дома, отношения с противоположным полом (секс), качество семейных отношений, свободное время после занятий, прогулки с друзьями, употребление алкоголя в рабочий день, употребление

---

<sup>51</sup> Министерство здравоохранения Российской Федерации. ПРИКАЗ от 1 октября 2020 года N 1053н «Об утверждении стандартов медицинской помощи взрослым при сахарном диабете 1 типа» [Электронный ресурс]. – URL: <https://docs.cntd.ru/document/566406946>

<sup>52</sup> Общество нефрологов, врачей диализа и трансплантологов Казахстана. Международные клинические рекомендации [Электронный ресурс]. – URL: <https://nephro.kz/ru/international-clinical-guidelines.html>. Об утверждении Стандарта организации оказания эндокринологической помощи в Республике Казахстан. Приказ Министра здравоохранения и социального развития Республики Казахстан от 6 июня 2016 года № 478. Зарегистрирован в Министерстве юстиции Республики Казахстан 5 июля 2016 года № 13880 [Электронный ресурс]. – URL: <https://adilet.zan.kz/rus/docs/V1600013880>

алкоголя по выходным, количество прогулов, текущее состояние здоровья (`health_level`).

По рисунку 1.21 можно проследить категорию уровня здоровья среди рассматриваемых студентов. Такая визуализация позволяет отслеживать и оценивать уровень здоровья учащихся по результатам каждого сезонного осмотра.

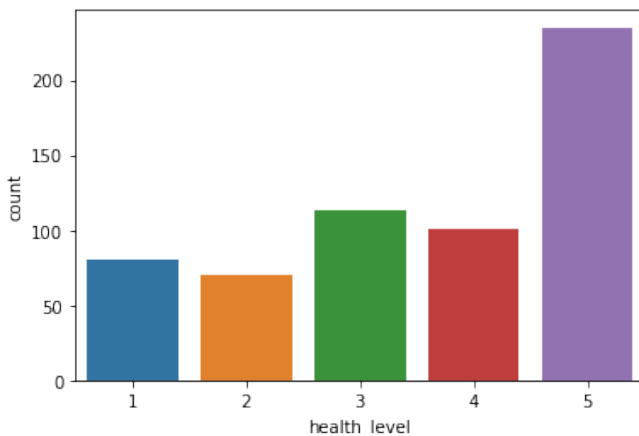


Рисунок 1.21 - Распределение уровня здоровья по учащимся

С помощью матрицы корреляций (Рис. 1.22) можно оценивать взаимосвязь (влияние) выбранных параметров и состояние здоровья студентов. Результаты анализа позволяют отметить взаимосвязь между потреблением алкоголя студентом и категорией здоровья. Такие исследования позволяют изучать здоровье студента в демографическом и социологическом контексте, не вдаваясь в их независимые патологические характеристики. Полученные результаты могут быть использованы при эффективном распределении учебной и методической нагрузки, учитывая, что такие параметры, как недельная нагрузка студентов (`study_time`) и свободное время, влияют на здоровье студента.

Использование таких алгоритмов, как Random Forest Classifier, SVM, Gradient Boosting Classifier, позволяет оценить зависимость `health_level` от других параметров. Например, можно

оценить профиль студентов, часто отсутствующих на занятиях, и попытаться выяснить причины, которые приводят к большому количеству пропусков, чтобы спрогнозировать дальнейшее поведение студента. Становится возможным изучить зависимость отсутствия студента на занятиях от параметров окружающей среды, которые отличаются от физиологического контекста.

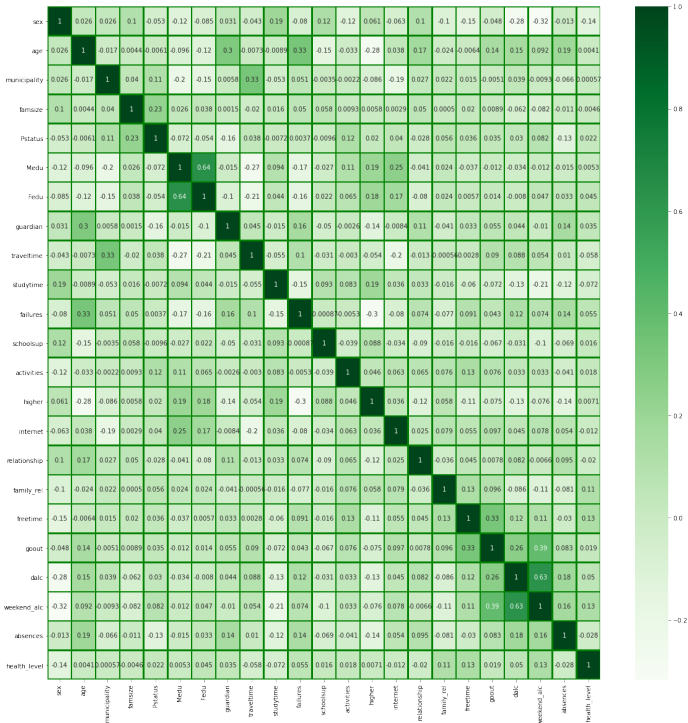


Рисунок 1.22 - Матрица корреляции

Алгоритм random forest использует множество решающих деревьев. Деревья строятся независимо друг от друга. При классификации документу присваивается класс, который определил наибольшее количество деревьев. В нашем случае выборка относительно невелика. Поэтому продвинутые алгоритмы на нем

не обладают высокой точностью. В результате было установлено, что алгоритм принятия решений по классификатору “случайный лес” является интересным выбором в качестве окончательной модели для использования в будущих наблюдениях.

Результаты на рисунке 1.23 показывают такие параметры, как размер семьи, потребление алкоголя, возраст, время еженедельных занятий, параметры, наиболее влияющие на состояние здоровья.

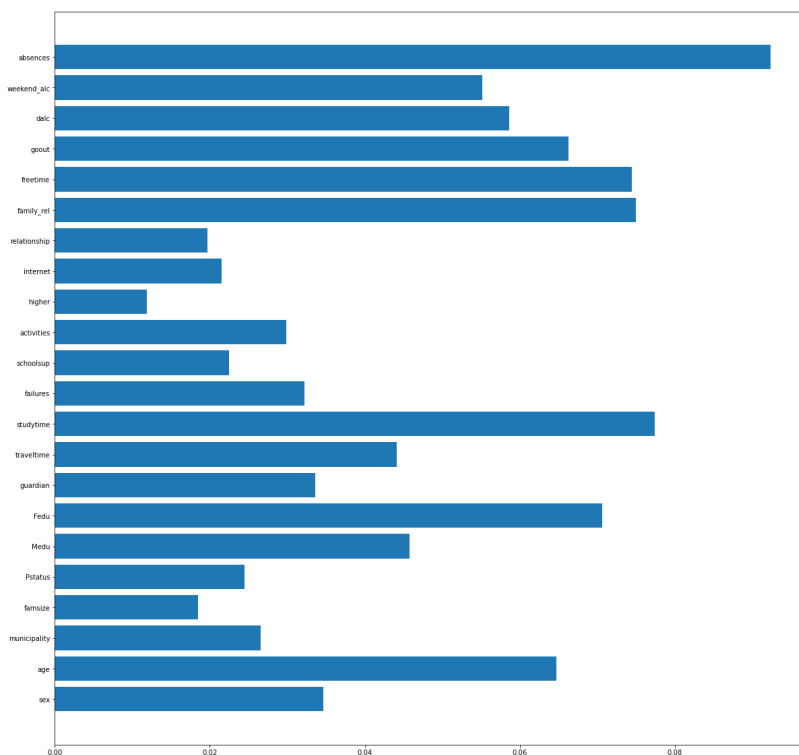


Рисунок 1.23 - Классификатор Random Forest:  
Зависимость/взаимосвязь уровня здоровья от других параметров

Оценка модели классификатора “случайный лес”:

– Количество объектов - 600, в том числе для обучения - 420 и тестирования - 180;

- Оценка точности обучающих данных составляет: 0,65;
- Оценка точности тестовых данных составляет: 0,57.

Такие анализы дают возможность изучить состояние здоровья студента не только в контексте медицинских данных. Это позволяет выявить связи между состоянием здоровья и социально-демографическими параметрами студента, оптимизировать принятие решений в задачах по организационным и административным вопросам университета, таким как распределение нагрузки, психологическая и социальная поддержка студента и т.д. Конечно, эти результаты не являются окончательными, и в будущем главной целью является расширение исследовательской базы и совершенствование аппарата искусственного интеллекта.

Таким образом, ценность предлагаемого подхода заключается в сборе на регулярной и непрерывной основе целенаправленной информации о состоянии здоровья студенческой молодежи. Кроме того, студенты, как и другие категории обучающихся, принадлежат к группам организованного населения, что удобно для успешного осуществления целенаправленных оздоровительных и профилактических мероприятий. Цель может быть достигнута путем накопления данных различных университетов на базе информационного портала КазНУ или путем предоставления технологии другим образовательным организациям, в том числе на коммерческой основе.

### ***1.2.7 Обеспечение безопасности и целостности данных***

Для обеспечения безопасности и целостности данных в информационно-аналитической системе «Цифровой профиль здоровья студента» используется способ разделения клиентского программного интерфейса от внутренней реализации – API Gateway (Рис. 1.24). Авторизация реализуется с использованием токенов JWT, что позволяет осуществлять контроль доступа в API Gateway. С помощью механизма аутентификации на основе токенов необходимость в учетных данных клиента заменяется токеном, который обеспечивает эффективное сохранение конфиденциальности клиента.

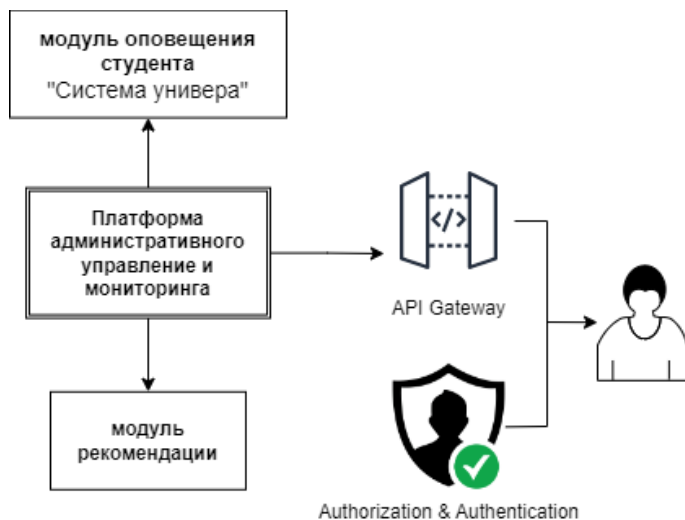


Рисунок 1.24 - Доступ к приложению

Наиболее важными причинами, по которым используются API-шлюзы, являются:

- Централизация и упрощение общего управления API и контроля;
- Повышение производительности и пользовательского опыта;
- Повышение и централизация безопасности, авторизации и аутентификации, а также управления токенами;
- Унификация ведения журнала, сбора метрик и наблюдения за состоянием системы с помощью расширенной панели мониторинга.

Система охватывает следующие аспекты безопасности: доступ на основе ролей, механизмы сетевой безопасности, шифрование данных, цифровая подпись и мониторинг доступа. Он использует доступ безопасности для распределения ролей и привилегий, которые подробно определены на уровне каждого действия. Механизм аутентификации позволяет каждому пользователю получать доступ к системе в пределах разрешенных ему привилегий. Каждый пользователь может получить к нему доступ,

войдя в веб-приложение или через веб-службу из приложения. А показатели для каждого студента интегрируются в личный кабинет студента.

### **1.3. Структура интеллектуальной информационно-аналитической системы и прототип веб-приложения**

Цифровое здоровье – это «широкий зонтичный термин, охватывающий электронное здравоохранение, а также развивающиеся области передовых компьютерных наук в области «больших данных», геномики и искусственного интеллекта»<sup>53</sup>. Цифровое здравоохранение ориентировано на потребителя благодаря использованию смартфонов и носимых датчиков, подключенных через Интернет вещей. Электронное здравоохранение было определено как «рентабельное и безопасное использование ИКТ для поддержки здоровья и связанных со здоровьем областей, включая медицинские услуги, медицинское наблюдение и медицинскую литературу, образование, знания и исследования»<sup>54</sup>.

Разрабатываемая интеллектуальная информационно-аналитическая система “Цифровой профиль здоровья студента” является реализацией подходов цифрового здоровья в РК. В данном разделе будут представлены структура интеллектуальной информационно-аналитической системы и прототип веб-приложения, описаны использованные технологии.

Одним из наиболее важных факторов при разработке приложений является выбор платформы, языка программирования и среды разработки. Для возможности разрабатывать и устанавливать мобильные приложения для Android и iOS на одной платформе был использован React.

React Native - это кроссплатформенный фреймворк с открытым исходным кодом, разработанный компанией Facebook для разработки нативных мобильных и настольных приложений на

<sup>53</sup> [Электронный ресурс]. – URL: [https://apps.who.int/gb/e/e\\_eb142.html](https://apps.who.int/gb/e/e_eb142.html)

<sup>54</sup> [Электронный ресурс]. – URL : [https://apps.who.int/gb/or/e/e\\_waha58r1.html](https://apps.who.int/gb/or/e/e_waha58r1.html)

JavaScript и TypeScript. iOS, Android, Windows и macOS могут поддерживаться для разработки кросс-платформенных приложений.

Преимуществами React являются:

- кроссплатформенность.
- простота и удобство разработки.
- большое количество документации.

Первоначальное знакомство пользователя с интеллектуальной информационно-аналитической системой “Цифровой профиль здоровья студента” происходит на главной странице, где кратко и понятно описаны основные цели и возможности системы. Доступ к основному интерфейсу имеется у всех пользователей системы, в целях конфиденциальности личные профайлы студентов зашифрованы. Последовательность действий и логика архитектуры информационно-аналитической системы представлена на рисунке 1.25.

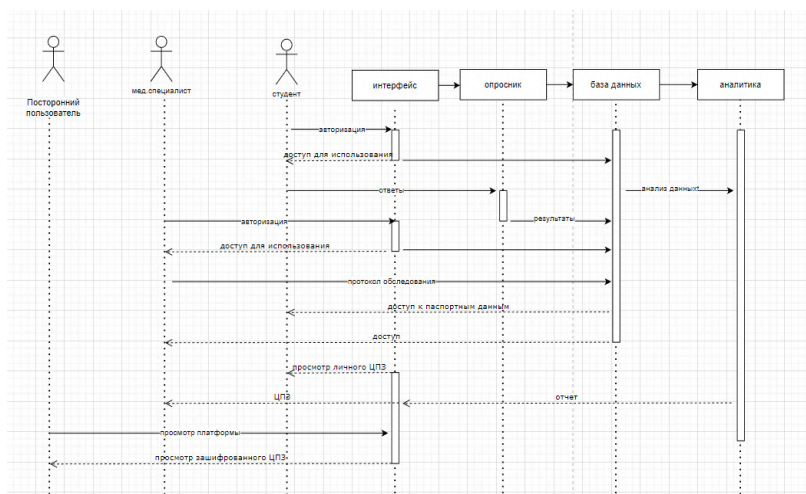


Рисунок 1.25 - Диаграмма последовательности действий

На главной странице находится основная информация (Рис. 1.26). Поясняется, что эта платформа для создания цифрового профиля здоровья студентов, здесь указаны доступные методы сбора данных и аналитические отчеты.





Рисунок 1.26 - Главная страница

Аналитические отчеты представлены в открытом доступе для просмотра, конфиденциальность личных данных студента обеспечены. Полученный html-код внедрения после публикации отчета в Power BI был размещен в системе. Сформированные отчеты визуализируют корреляционные связи, тренды, прогнозы по данным.

В разработанном прототипе пользователями системы являются:

- медицинский работник;
- администратор системы;
- студент.

Как показано на рисунке 1.27, студент или врач входят в приложение, регистрируясь и используя свои собственные логины.

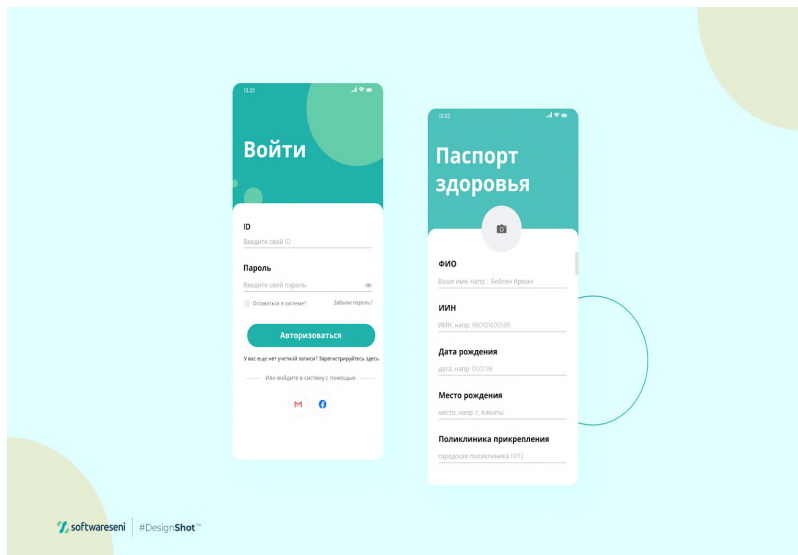


Рисунок 1.27 - Аутентификация пользователя

После регистрации и входа в систему для медицинского персонала и пациента открываются два разных представления с соответствующим функционалом. Для медицинского персонала создается пара логин - пароль на каждого работника. База данных, предназначенная для ввода данных пациентов (студентов), разработана по концепции электронной медицинской карты, которая содержит дополнительную информацию о пациенте (студенте), в том числе фото, ранние заболевания, аллергические реакции, различные противопоказания, а также историю развития заболевания.

Профиль студента содержит паспортные данные, ИИН, место рождения, дату рождения, полный возраст, пол, физикометрические данные, прививочный паспорт, результаты медицинских осмотров, анализы (Рис. 1.28).

**ЛИЧНЫЕ ДАННЫЕ**

**ОПРОСНИК**

- социальная поддержка
- выявление триггеров
- опрос о стрессе
- ИБС

**АНАЛИТИКА**

**МЕДСМОТР**

- профилактический
- периодический
- внеочередной

**Общая информация**

- Фамилия: Асылбек
- Имя: Асхат
- Отчество: Жаркынович
- ИИН: 020411400586
- Дата рождения: 11.04.2002
- Место рождения: г. Орал
- Национальность: казак
- Адрес постоянного проживания: г. Алматы
- Адрес проживания в г. Алматы: аль-Фараби 71/9
- Поликлиника прикрепления: "Smart health university city"

**Анкета для определения риска сахарного диабета 2-го типа**

**Открыть чат**

**РЕДАКТИРОВАТЬ** **СОХРАНИТЬ**

Рисунок 1.28 - Личный профиль студента

У каждого медицинского работника рассмотрена возможность вводить и редактировать данные по итогам обследований по своей компетенции. По вводимым данным указаны интерпретации, ввод числовых данных и меню выборки для текстовых описаний (Рис. 1.29-1.35).

**Кардиолог**

**ЭКГ**

**тип** **тип ЭКГ** **интерпретация**

- стандартная
- не стандартная
- другое
- нормально
- патологич.
- другое

**интерпретация**

- нормально
- патологич.
- другое

**Измерение АД**

**тип** **тип измерения** **систолическая**

- диастолическая
- систолическая

**Пульс**

**интерпретация**

- нормально
- патологич.
- бранинадия

**Аускультация сердца**

**интерпретация** **тип аускультации**

- нормально
- патологич.
- диастолическая
- систолическая
- патологич.
- другое
- нормально
- патологич.
- другое

Рисунок 1.29 - Ввод данных Кардиолога

⚙️

## Невролог

### Рефлексы

коленный		карпорадиальный	
<b>левый</b>	<b>правый</b>	<b>левый</b>	<b>правый</b>
<input type="radio"/> норма <input type="radio"/> повышены <input type="radio"/> снижены <input type="radio"/> отсутствуют	<input type="radio"/> норма <input type="radio"/> повышены <input type="radio"/> снижены <input type="radio"/> отсутствуют	<input type="radio"/> норма <input type="radio"/> повышены <input type="radio"/> снижены <input type="radio"/> отсутствуют	<input type="radio"/> норма <input type="radio"/> повышены <input type="radio"/> снижены <input type="radio"/> отсутствуют

---

### Чувствительность

<b>болевая</b>	<b>температурная</b>	<b>тактильная</b>
<input type="radio"/> норма <input type="radio"/> гипостезия <input type="radio"/> гиперестезия <input type="radio"/> анестезия	<input type="radio"/> норма <input type="radio"/> повышены <input type="radio"/> снижены <input type="radio"/> отсутствуют	<input type="radio"/> присутствует <input type="radio"/> отсутствуют

Рисунок 1.30 - Ввод данных Невролога

### Высшие корковые функции

<b>гипсис</b>	<b>праксис р/а/л/р</b>	<b>ритм</b>	<b>письмо</b>
<input type="radio"/> острое амнезия <input type="radio"/> эпилепсия <input type="radio"/> астереогноз <input type="radio"/> другое	<input type="radio"/> апраксия <input type="radio"/> диспраксия <input type="radio"/> другое	<input type="radio"/> афазия мотор <input type="radio"/> афазия сенс <input type="radio"/> афазия сенс/мотор <input type="radio"/> афазия Брока <input type="radio"/> другое	<input type="radio"/> аграфия <input type="radio"/> алексия <input type="radio"/> дислексия <input type="radio"/> другое

---

### Черепно-мозговые нервы

<b>1nч</b>	<b>3nч 4nч 6nч</b>	<b>5nч</b>	<b>7nч</b>
<input type="radio"/> гипосмия <input type="radio"/> anosmia <input type="radio"/> норма	<input type="radio"/> ptosis <input type="radio"/> ptosis <input type="radio"/> миоз <input type="radio"/> диплопия <input type="radio"/> нарушение конвергенции <input type="radio"/> косоглазие <input type="radio"/> норма <input type="radio"/> анисокория <input type="radio"/> норма	<input type="radio"/> болезненность <input type="radio"/> тонус выходящих ветвей <input type="radio"/> норма	<input type="radio"/> гемипарез/паралич <input type="radio"/> лагофтальм <input type="radio"/> норма
<b>8nч</b>	<b>9nч 10nч</b>	<b>11nч</b>	<b>12nч</b>
<input type="radio"/> вестибулярная атакия <input type="radio"/> интенционный тремор <input type="radio"/> неустойчивость в пр <input type="radio"/> шаткость <input type="radio"/> миастия <input type="radio"/> норма	<input type="radio"/> дисфония <input type="radio"/> паралич <input type="radio"/> жал/отсут/никс <input type="radio"/> глоточного рефлекс <input type="radio"/> дисфония <input type="radio"/> норма	<input type="radio"/> нарушение поворота головы <input type="radio"/> вверх и влево/ вправо <input type="radio"/> норма	<input type="radio"/> глоссолатия <input type="radio"/> парез <input type="radio"/> односторонний <input type="radio"/> дизартрия <input type="radio"/> норма

Заключение:


✎ РЕДАКТИРОВАТЬ
СОХРАНИТЬ

Рисунок 1.31 - Ввод данных Невролога, продолжение

earning    ↻    Статья 'Нейросете...    🇬🇧 English Test for Ex...    📊 Исследование мат...

СПЕЦИАЛИСТЫ    ОПРОСНИК    АНАЛИТИКА

---

ОВ     **Офтальмолог**

**Авторефрактометрия**

**МИОПИЯ**

- Слабой степени от - 0,25 до - 3,0 Д
- Средней степени от - 3,25 до - 6,0 Д
- Высокой степени от - 6,25 и выше

**АСТИГМАТИЗМ**

- Простой миопический
- Простой гиперметропический
- Сложный миопический
- Сложный гиперметропический
- Смешанный

**ГИПЕРМЕТРОПИЯ**       **НОРМА**

---

**Измерение внутриглазного давления**

**НОРМА**

**ГИПЕРТЕНЗИЯ**

- 21 и выше

**ГИПОТОНИЯ**

- 10 и ниже

---

**Цветовосприятие**

**ВСЕ ЦВЕТА РАЗЛИЧАЕТ**

**ЦВЕТАНОМАЛИЯ**

- Дальтонизм
- Протаномалия
- Тританомалия
- Дейтераномалия

---

**Визиометрия**

**ПОНИЖЕННОЕ ЗРЕНИЕ**

**НОРМАЛЬНОЕ ЗРЕНИЕ**

- Прочитал 10 рядов без ошибок (с 8 ряда 2 ошибки допустимы)

Рисунок 1.32 - Ввод данных Офтальмолога

В данном разделе приведены примеры входных данных. Функционал портала, предназначенный для медицинских специалистов, дает возможность автоматизировать процесс протоколирования обследования. Все возможные варианты интерпретации

по медицинским показателям, опросники, методы развернутого обследования для получения полного анамнеза были подобраны и составлены специалистами Факультета медицины и здравоохранения КазНУ им. аль-Фараби.

Рисунок 1.33 - Ввод данных - Офтальмолога, продолжение

Список студентов  
АНАЛИТИКА

### Физиометрические показатели здоровья

**Длина тела**

Заполнить  (конец, начало, дата, время, источник, состояние)

**Вес**

Заполнить  (конец, начало, дата, время, источник, состояние)

**Индекс массы тела (ИМТ)**

Заполнить  (конец, начало, дата, время, источник, состояние)

**Объем талии (ОТ)**

Заполнить  (конец, начало, дата, время, источник, состояние)

Рисунок 1.34 - Ввод данных - Физиометрические показатели здоровья

Список студентов  
АНАЛИТИКА

### Прививочный паспорт

#### Туберкулез

1-ая		2-ая	
сроки	реакция	сроки	реакция
<input type="radio"/> не проводилась	<input type="radio"/> нет	<input type="radio"/> не проводилась	<input type="radio"/> нет
<input type="radio"/> своевременно	<input type="radio"/> слабая	<input type="radio"/> своевременно	<input type="radio"/> слабая
<input type="radio"/> отсроченно	<input type="radio"/> средняя	<input type="radio"/> отсроченно	<input type="radio"/> средняя
<input type="radio"/> нет данных	<input type="radio"/> высокая	<input type="radio"/> нет данных	<input type="radio"/> высокая

#### Вирусный гепатит «В»

1-ая		2-ая		3-ая	
сроки	реакция	сроки	реакция	сроки	реакция
<input type="radio"/> не проводилась	<input type="radio"/> нет	<input type="radio"/> не проводилась	<input type="radio"/> нет	<input type="radio"/> не проводилась	<input type="radio"/> нет
<input type="radio"/> своевременно	<input type="radio"/> слабая	<input type="radio"/> своевременно	<input type="radio"/> слабая	<input type="radio"/> своевременно	<input type="radio"/> слабая
<input type="radio"/> отсроченно	<input type="radio"/> средняя	<input type="radio"/> отсроченно	<input type="radio"/> средняя	<input type="radio"/> отсроченно	<input type="radio"/> средняя
<input type="radio"/> нет данных	<input type="radio"/> высокая	<input type="radio"/> нет данных	<input type="radio"/> высокая	<input type="radio"/> нет данных	<input type="radio"/> высокая

Рисунок 1.35 -Прививочный паспорт студента

#### 1.4. Список статей, опубликованных в рамках исследования:

1. Mansurova M., Zubairova M., Kadyrbek N., Tyulepberdinova G., Sarsembayeva T. Data Analysis for The Student Health Digital Profile. 2021 16th International Conference on Electronics Computer and Computation (ICECCO) | 978-1-6654-0945-2/21/\$31.00 ©2021 IEEE | DOI: 10.1109/ICECCO53203.2021.9663804

2. Zholdas N., Mansurova M., Postolache O., Kalimoldayev M., Sarsembayeva T. // A Personalized mHealth Monitoring System for Children and Adolescents with T1 Diabetes by Utilizing IoT Sensors and Assessing Physical Activities//. International Journal of Computers, Communications and Control this link is disabled, 2022, 17(3), 4558.

3. Zholdas N., Mansurova M., Sarsembayev M., Shomanov A., Sarsembayeva T. Application of mHealth Technologies to Improve Self-Control of Children and Adolescents with Type 1 Diabetes // 2022 IEEE International Symposium on Medical Measurements and Applications, MeMeA 2022 - Conference Proceedings, 2022

4. Tasmurzayev N., Amangeldy B., Baigarayeva Z., Resnik B., Amirkhanova G. Improvement of HVAC System Using the Intelligent Control System // ENERGYCON 2022 - 2022 IEEE 7th International Energy Conference, Proceedings, 2022.

5. Zholdas N., Postolache O., Mansurova M. Health Monitoring System Using Internet of Things // SIST 2021 - 2021 IEEE International Conference on Smart Information Systems and Technologies, 2021, 9465928

6. Mansurova M.E., Barakhnin V.B., Tyulepberdinova G.A., Gusmanova F.R., Nurakhanova A.A. Comparative analysis of the classification of health status by artificial intelligence methods// Вестник КазНПУ им. Абая, серия «Физико-математические науки», №3(75), 2021 г.

7. Тюлепбердинова Г.А., Мансурова М.Е., Шілмағамбетова А. Методы и приемы используемые при анализе данных для оценки состояния здоровья студентов // VI Международная научно-практическая конференция «Информатика и прикладная математика. 29 сентября - 2 октября 2021 г., Алматы, Казахстан.



8. Мансурова М.Е., Тюлепбердинова Г.А., Сулеймен О.Д. Анализ информационно-аналитических систем, используемых в целях оценки состояния здоровья подростков // Материалы Международной научной конференции в области информационных технологий, посвященной 75-летию профессора У.А. Тукеева. Алматы, 8 октября 2021 года: – Алматы: Қазақ университеті, 2021. 23-28 стр.

9. Сулеймен О., Тюлепбердинова Г. // Региональная медицинская информационно-аналитическая система Материалы международной конференции студентов и молодых учёных «Фараби әлемі» Алматы, Казахстан, 6-8 апреля 2022 года, 14-15 стр.

## **2. АВТОМАТИЗИРОВАННАЯ СИСТЕМА ПО УЧЕТУ ПУБЛИКАЦИОННОЙ АКТИВНОСТИ УЧЕНЫХ И ОРГАНИЗАЦИЙ КАЗАХСТАНА МЕТОДОМ БИБЛИОМЕТРИЧЕСКОГО АНАЛИЗА НА ОСНОВЕ ДАННЫХ API SCORUS**

*Смагулов К.Е., Ұзақбай С.Ә., Омиралиева Г.К.,  
Кистаубаев Е.Б., Дусекеев Р.М.*

В данной главе приводится анализ существующих библиометрических методов оценки научной производительности ученых и организаций на основе публикационной активности в рейтинговых изданиях, индексируемых в международных базах данных. Рассмотрены действующие в различных странах методологии и системы оценки эффективности деятельности ученых и университетов на основе количественных и качественных показателей, оценивающих вклад ученого, организации, страны в написание научной статьи, цитируемость публикаций, международная коллаборация и др.

Приводится информация об авторской автоматизированной системе по оценке публикационной активности авторов и организаций Республики Казахстан на основе данных из базы Scopus и с применением метода анализа с учетом требований, стоящих перед казахстанскими учеными в настоящее время и соответствующих нормативно-правовым актам Республики.

В заключительном подразделе приведены техническое описание системы, а также методология ранжирования ученых и организаций Казахстана на основе формирования их рейтинга по количественным и качественным показателям публикационной активности.

*Данная глава монографии подготовлена в рамках выполнения научного проекта на тему АР08051974 «Оценка интеллектуального капитала вузов Казахстана на основе наукометрического анализа и пути их трансформации в исследовательские универ-*

ситеты» грантового финансирования молодых ученых Комитета науки МОН РК на 2020-2022 гг.

## **2.1. Анализ библиометрических методов оценки эффективности публикационной активности авторов и организаций**

Оценка научной производительности позволяет оценить и измерить вклад ученых и организаций, занимающихся научно-исследовательской деятельностью. Одним из основных показателей результативности научных исследований является публикационная активность в рейтинговых и международных изданиях.

В практике ведущих государств мира основной методикой оценки интеллектуального капитала, эффективности ученых на основе публикаций является наукометрический анализ, в частности библиометрия<sup>55</sup>.

Библиометрия – многогранное направление, включающее структурную, прогностическую, динамическую и оценочную наукометрию в развитии отрасли науки с использованием численных и статистических методов<sup>56</sup>. Библиометрический метод предлагает полезный инструмент перехода с микроуровня (ученые и институты) на макроуровень (национальный и глобальный)<sup>57</sup>. Так, количественная оценка, анализ и сравнение научных результатов исследователей важны для правительства, институтов и университетов<sup>58</sup>. Так, в некоторых странах финансирование научных проектов зависит от показателя публикационной активности ученого, выступающего в качестве руководителя.

<sup>55</sup> Hicks, D., Wouters, P., Waltman, L., De Rijcke, S., & Rafols, I. Bibliometrics: The Leiden manifesto for research metrics // *Nature*. – 2015. – Vol. 520(7548). – P.429-431.

<sup>56</sup> Guozhu M., Xi L., Hui bin D., Jian Z., Linyuan W. Way forward for alternative energy research: A bibliometric analysis during 1994–2013 // *Renewable and Sustainable Energy Reviews*. – 2015. – Vol. 48. – P. 276-286.

<sup>57</sup> Zhou F, Guo HC, Ho YS, Wu CZ. Scientometric analysis of geostatistics using multivariate methods // *Scientometrics* – 2007. – Vol.73(3). – P.265–79.

<sup>58</sup> Senanayake, U., Piraveenan, M., Zomaya, A. The pagerank-index: Going beyond citation counts in quantifying scientific impact of researchers // *PLoS ONE*. – 2015. – Vol.10(8). doi:10.1371/journal.pone.0134794

Подсчет с помощью библиометрических методов необходим, поскольку известно, что их производительность измеряется и приводится в различных рейтингах. Известным примером является CWTS (*Centrum voor Wetenschap en Technologische Studies*) Leiden Ranking<sup>59</sup> – всемирная рейтинговая система университетов, основанная на научном вкладе и влиянии. В качестве источника данных публикаций и цитирования используется библиографическая база данных Web of Science компании Clarivate Analytics.

Лейденский рейтинг включает в себя три показателя цитируемости статей авторов университета<sup>60</sup>:

– MCS (*Mean citation score*) вычисляет среднее количество цитируемости университетских публикаций, то есть общее количество цитируемости делится на общее количество публикаций.

– MNCS (*Mean normalized citation score*) влияние цитирования (количество цитирования на статью), нормализованное по теме, году и типу документа. Значение MNCS можно интерпретировать как соответствующее мировому уровню, если оно равно единице.

– PP<sub>top 10%</sub> (*Proportion top 10% publications*) доля публикаций в топ-10% на основе цитирования по категориям, годам и типам документов. То есть публикации университета в журналах с высоким процентилем показывают, насколько они представлены в самых цитируемых журналах мира.

Тем не менее, Лютт Лейдесдорф (2010) в своей статье критиковал метод, используемый в рейтинге Лейдена для нормализации воздействия цитирования (MNCS), как основанный на классификации предметных категорий ISI (Institute for Scientific Information), используемых в Web of Science, так как последний разработан для поиска информации, а не для наукометрической оценки<sup>61</sup>.

<sup>59</sup> “Data collection”, CWTS Leiden Ranking. Universiteit Leiden Centre for Science and Technology Studies [Электронный ресурс]. URL: <https://www.leidenranking.com/>

<sup>60</sup> Waltman, L. R., Calero Medina, C. M., Kosten, J., Noyons, E. C. M., Tijssen, R. J. W., Eck, N. J. P. van, ... Wouters, P. (2012). The Leiden Ranking 2011/2012: Data collection, indicators, and interpretation. Centre for Science and Technology Studies, Leiden University. Retrieved from <https://hdl.handle.net/1887/19353>.

<sup>61</sup> Leydesdorff L., Opthof T., Normalization, CWTS indicators, and the Leiden Rankings: Differences in citation behavior at the level of fields. – 2010.

Выбор метода подсчета важен для большинства библиометрических анализов, которые оценивают авторов, учреждения, страны и др. Изучение выборки из 99 библиометрических исследований показывает<sup>62</sup>, что для двух третей исследований выбор метода подсчета может повлиять на результаты.

Библиометрические показатели можно разделить на три основные группы<sup>63</sup>:

- Количественные показатели или показатели производительности, например, количество публикаций.
- Хирша (h-index).
- Структурные показатели позволяют оценить связи сотрудничества (коллораации), например, соавторов из разных областей, учреждений или стран.

Во многих странах мира разработаны, утверждены и широко используются системы оценки эффективности деятельности ученых и университетов на основе их публикационной активности, цитируемости и коллаборации с точки зрения библиометрии. Данные системы играют ключевую роль, как в распределении финансирования, так и карьерного роста ученых<sup>64</sup>. В свою очередь, научные публикации являются прологом патентов, то есть новых продуктов, технологий, на основе которых университет обеспечивает свой успех<sup>65</sup>.

При оценке научной производительности ученых и организаций на основе библиометрии основной единицей измерения являются научные публикации, в первую очередь опубликованные в изданиях, входящих в международные базы данных

---

<sup>62</sup> Marianne G. Counting methods introduced into the bibliometric research literature 1970–2018: A review // *Quantitative Science Studies*. – 2021. – Vol.2(3). – P.932-375. doi.org/10.1162/qss\_a\_00141.

<sup>63</sup> Nikolic, S., Penca, V., Ivanovic, D., Surla, D. Konjovic, Z. Storing of Bibliometric Indicators in CERIF Data Model [Электронный ресурс]. – URL: [https://www.researchgate.net/publication/271015214\\_Storing\\_Of\\_Bibliometric\\_Indicators\\_In\\_Cerif\\_Data\\_Model](https://www.researchgate.net/publication/271015214_Storing_Of_Bibliometric_Indicators_In_Cerif_Data_Model).

<sup>64</sup> Abbott A, Cyranoski D, Jones N, et al. Metrics: Do metrics matter? // *Nature*. - 2010. – Vol. 465, No 2. - P. 860-862.

<sup>65</sup> Kalemis, K., Saba, R. & Elpida, A. The Value of Measuring Intellectual Capital (IC) In Higher Education – A New Challenge of Our Days [Электронный ресурс]. – URL: [https://conferences.ionio.gr/icil2012/download.php?f=papers/186-bou-saba-full\\_text-en-v001.pdf](https://conferences.ionio.gr/icil2012/download.php?f=papers/186-bou-saba-full_text-en-v001.pdf)

Scopus и/или Web of Science (WoS). Использование именно этих баз данных обусловлено тем, что они имеют широкий функционал по наукометрическим показателям изданий, публикаций, авторов, организаций и стран. Каждая из этих баз имеет свои особенности, такие как глубина охвата документопотока и репертуар индексируемых изданий. Поэтому представленные в них данные также различаются<sup>66</sup>.

Базы данных Scopus и Web of Science индексируют журналы с высоким рейтингом, так как существует довольно сложная и тщательная система отбора научного издания для индексации. Обе базы данных имеют обширный функционал, позволяющий анализировать данные на основе различных параметров: количество цитирований публикаций, рейтинг журнала, индекс Хирша и др. Однако, в базе данных Scopus примерно в два раза больше индексируемых изданий, нежели в WoS. Соответственно, число цитирований публикаций, как правило, выше у Scopus, чем в WoS<sup>67</sup>.

Тем не менее, в WoS, так и в Scopus журналы по общественным и гуманитарным наукам не широко представлены, а также в обеих базах абсолютное большинство изданий англоязычные<sup>68</sup>.

С ростом числа публикаций становится очевидным, что количество публикаций не может дать конкретной оценки исследованию, учреждению или ученому<sup>69</sup>. Качество научной статьи

---

<sup>66</sup> Meho, Lokman I., and Cassidy R. Sugimoto. «Assessing the Scholarly Impact of Information Studies: A Tale of Two Citation Databases – Scopus and Web of Science» // *Journal of the American Society for Information Science and Technology*. – 2009. – Vol.60(12). – P.2499–2508.

<sup>67</sup> Haddow, G., Genoni, P. Citation analysis and peer ranking of Australian social science journals // *Scientometrics*. – 2010. – Vol.85(2), – P.471-487.; Kulkarni, A. V., Aziz, B., Shams, I., & Busse, J. W. Comparisons of citations in Web of Science, Scopus, and Google Scholar for articles published in general medical journals // *JAMA*. – 2009. – Vol.302(10), – P.1092-1096.; Torres-Salinas, D., Lopez-Cózar, E. D., Jimenez-Contreras, E. Ranking of departments and researchers within a university using two different databases: Web of Science versus Scopus // *Scientometrics*. – 2009. – Vol.80(3), – P.761-774.

<sup>68</sup> Mongeon, P., & Paul-Hus, A. (2016). The journal coverage of Web of Science and Scopus: A comparative analysis // *Scientometrics*. – 2016. – Vol.106(1), – P. 213-228.

<sup>69</sup> Abramo G. Revisiting the scientometric conceptualization of impact and its measurement // *Journal of Informetrics*. – 2018. – Vol.12(3), – P.590–597.

или отдельного ученого в рейтинговых изданиях можно оценить по количеству цитирований в трудах других авторов<sup>70</sup>. Количество публикаций ученого отражает объем проделанной им работы, а индекс цитируемости гораздо ближе к качественному показателю этой работы, так как связан с такими факторами, как оригинальность, интерес научного сообщества и актуальность конкретных тем исследований. Статьи с высокой цитируемостью признаются доказательством исключительной исследовательской деятельности<sup>71</sup>.

Показатели журналов, индексируемых в базах данных Scopus и WoS, основываются на цитируемости статей. Так, в базе данных WoS применяется показатель импакт-фактора журнала, а в базе Scopus – CiteScore. Обе метрики высчитываются соотношением количества цитирований на количество публикаций по каждому журналу.

Таким образом, чем выше показатель импакт-фактора или CiteScore журнала, то, вероятнее, статья, которая будет опубликована в данном журнале, имеет больше шансов быть процитированной. И наоборот: чем меньше показатели импакт-фактора и CiteScore издания, то высока вероятность того, что цитируемость статьи будет низкой.

Отдельные европейские университеты используют показатель импакт-фактора журнала вместо индекса цитируемости как показателя качества публикаций в своих рейтингах.

Необходимо также отметить, что имеются авторы, которые стремятся увеличить свои показатели цитируемости посредством самоцитирования, то есть в своих статьях ссылаются на ранее опубликованные свои труды. Некоторые ученые, а также организации, применяющие оценки эффективности научной производительности своих сотрудников на основе библиоме-

---

<sup>70</sup> Gupta, H.M., Campanha, J.R., Pesce, R. Power-law distributions for the citation index of scientific publications and scientists // *Brazilian Journal of Physics*. – 2005. – Vol.35(4 A), – P.981-986. doi:10.1590/S0103-9733200500060001.

<sup>71</sup> Dag W. Aksnes, Liv Langfeldt, Paul Wouters. Citations, Citation Indicators, and Research Quality: An Overview of Basic Concepts and Theories // *SAGE Open*. – 2019. – Vol.1(7). DOI: 10.1177/2158244019829575.

трии, считают, что самоцитирование является нежелательным фактором при анализе цитирования автора и что самоцитирование не должно применяться в расчет<sup>72</sup>. С другой стороны, отдельные ученые подвергают сомнению негативный эффект самоцитирования<sup>73</sup>.

Тем не менее, очевидно, что самоцитирование существенно влияет на индекс Хирша автора. Данный индекс позволяет определить уровень цитируемости автора, посредством формулы расчета h-индекса:  $n$  количество статей, процитированных не менее  $n$  количество раз каждая<sup>74</sup>. Следует также иметь в виду, что h-индекс изменяется в соответствии с научной базой данных, использованной для его расчета<sup>75</sup>. Соответственно, один и тот же автор имеет разный показатель индекса Хирша согласно базе данных Scopus и WoS. Тем не менее, данная метрика имеет ряд недостатков, один из которых то, что им можно манипулировать посредством самоцитирований, а также он не отражает реальный вклад ученого в публикацию<sup>76</sup>.

В целом, выявление вклада каждого автора в написание статьи с числом авторов более 1 является одной из актуальных в современной библиометрии. Существуют различные подходы по учету вклада между соавторами одной публикации. Важность применения этих методов имеет значение для всех рейтингов, основанных на количестве публикаций и цитирований<sup>77</sup>.

---

<sup>72</sup> Aksnes, D. W. A macro study of self-citation // *Scientometrics*. – 2003. – Vol. 56(2). – P. 235-246.

<sup>73</sup> Fowler, J. H., Aksnes, D. W. (2007). Does self-citation pay // *Scientometrics*. – 2007. – Vol.72(3), – P. 427–437. <https://doi.org/10.1007/s11192-007-1777-2>.

<sup>74</sup> Hirsch, J. E. An index to quantify an individual's scientific research output // *Proceedings of the National Academy of Sciences of the United States of America*. – 2005. - Vol. 702, No 46. – P. 16569-16572.

<sup>75</sup> Lippi G, Borghi L. A short story on how the H-index may change the fate of scientists and scientific publishing // *Clinical Chemistry and Laboratory Medicine (CCLM)*. – 2014. – Vol.52. – P.1-3. <https://doi.org/10.1515/cclm-2013-0715>.

<sup>76</sup> Lippi, G., Mattiuzzi, C. Scientist impact factor (SIF): A new metric for improving scientists' evaluation? // *Annals of Translational Medicine*. – 2017. – Vol.5(15). doi:10.21037/atm.2017.06.24.

<sup>77</sup> Gauffriau, M., Larsen, P. O. Counting methods are decisive for rankings based on publication and citation studies // *Scientometrics*. – 2005. – Vol. 64, No 1. - P. 85-93.



Наиболее объективной схемой распределения кредита (вклада) между соавторами является метод, когда сами авторы указывают внесенный вклад каждого при написании статьи<sup>78</sup>. Однако, эта практика применяется не во всех журналах и авторских коллективах. В связи с этим были разработаны методы определения вклада каждого автора в написание публикации, основными из которых являются:

- Полный подсчет,
- Прямой подсчет,
- Дробный подсчет,
- Взвешенный подсчет.

Под полным подсчетом (*whole counting*) понимается равномерное распределение по одному кредиту всем соавторам статьи. Данный метод не является сбалансированным и объективным методом оценки, так как имеет инфляционные погрешности и приносит пользу второстепенным авторам за счет первичных авторов<sup>79</sup>. Кроме того, индивидуальный вклад десяти соавторов в статью сильно отличается от того, если статья написана одним автором. В схеме полного подсчета нарушается базовое условие – сумма авторских кредитов равна 1 для каждой статьи, гарантирующее, что все документы имеют одинаковый вес независимо от количества соавторов<sup>80</sup>.

У полного подсчета существуют различные вариации, когда один авторский кредит распределяется между: авторами, организациями соавторов, стран соавторов<sup>81</sup>.

Наиболее распространенным методом учета вклада первого и корреспондентом автора называется прямым подсчетом (*straight*

---

<sup>78</sup> Vavrycuk V. Fair ranking of researchers and research teams [Электронный ресурс]. - 2018. – URL: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0195509>

<sup>79</sup> Hagen, N. T. Harmonic allocation of authorship credit: Source-level correction of bibliometric bias assures accurate publication and citation analysis [Электронный ресурс]. - 2008. – URL: doi:10.1371/journal.pone.0004021

<sup>80</sup> Vavrycuk V. Fair ranking of researchers and research teams [Электронный ресурс]. - 2018. – URL: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0195509>

<sup>81</sup> Gauffriau, M., Larsen, P. O., Maye, I., Roulin-Perriard, A., Von Ins, M. Publications, cooperation and productivity measures in scientific research // *Scientometrics*. – 2007. – Vol. 73, No 2. - P. 175-214.

*counting*). Данный метод состоит в том, чтобы предоставить полный кредит публикации только первому автору и не давать кредиты другим авторам<sup>82</sup>. Также данный метод применяется для учета кредита только корреспонденту автору<sup>83</sup>. Важность прямого подсчета заключается в том, что должность первого автора относится к «генератору идей» и сотрудникам, выполняющим основную часть работы, а последняя позиция (корреспондент автор) назначается общему руководителю рабочей группы<sup>84</sup>. Идея метода заключается только в том, чтобы дать кредит главному лидеру. Поэтому группа Scimago использовала термин «лидерство» для этого типа подсчета<sup>85</sup>.

Одним из распространенных методов распределения авторского кредита является дробный подсчет (*fractional counting*), который заключается в разделении 1 кредита между всеми соавторами статьи. Дробный подсчет может быть реализован разными способами, когда равная доля кредита распределяется между: авторами; организациями; странами<sup>86</sup>.

Однако, и дробный метод несовершенен, так как вклад каждого автора не является равноценным. Поэтому, в лите-

---

<sup>82</sup> Lindsey, D. Production and citation measures in the sociology of science: The problem of multiple authorship // *Social Studies of Science*. – 1980. – Vol. 10, No 2. – P. 145-162. Lange, L. L. Citation counts of multi-authored papers—First-named authors and further authors // *Scientometrics*. – 2001. – Vol. 52, No 3. – P. 457-470.

<sup>83</sup> Moya-Anegón, F., Guerrero-Bote, V.P., Bornmann, L., Moed, H.F. The research guarantors of scientific papers and the output counting: A promising new approach // *Scientometrics*. – 2013. – Vol. 97, No 2. – P. 421-434. Huang, M. H., Lin, C. S., & Chen, D. Z. Counting methods, country rank changes, and counting inflation in the assessment of national research productivity and impact // *Journal of the American Society for Information Science and Technology*. – 2011. – Vol. 62, No 12. – P. 2427-2436. Lin, C. S., Huang, M. H., Chen, D. Z. The influences of counting methods on university rankings based on paper count and citation count // *Journal of Informetrics*. – 2013. – Vol. 7, No 3. – P.611-621.

<sup>84</sup> Abramo, G., D'Angelo, C. A., Rosati, F. The importance of accounting for the number of co-authors and their order when assessing research performance at the individual level in the life sciences // *Journal of Informetrics*. – 2013. – Vol. 7, No 1. – P. 198-208.

<sup>85</sup> SCImago Research Group. Description of SCImago Journal Rank Indicator. Accessed 2012. URL: <http://www.scimagojr.com/SCImagoJournalRank.pdf> and <http://arxiv.org/ftp/arxiv/papers/0912/0912.4141.pdf>

<sup>86</sup> Waltman, L., & Van Eck, N. J. Field-normalized citation impact indicators and the choice of an appropriate counting method // *Journal of Informetrics*. – 2015. – Vol. 9, No 4. – P. 872-894.

ратуре были предложены различные подходы для предоставления кредита авторам публикации на основе их положения в списке авторов, причем первый автор обычно считается наиболее важным автором<sup>87</sup>. Данный метод не применяется в статьях, где авторы публикации располагаются в алфавитном порядке<sup>88</sup>.

Различие распределения авторского кредита по наиболее распространенным подходам указано в таблице 2.1.

Таблица 2.1

**Сравнительная таблица четырех различных методов подсчета авторского кредита<sup>89</sup>**

	Автор 1	Автор 2	Автор 3	Автор 4	Автор 5
Полный подсчет	1,00	1,00	1,00	1,00	1,00
Дробный подсчет на уровне автора	0,20	0,20	0,20	0,20	0,20
Подсчет первого автора	1,00	0,00	0,00	0,00	0,00
Подсчет корреспондент автора	0,00	0,00	0,00	1,00	0,00

С одной стороны, исследования ученых показали, что прямой подсчет и дробный подсчет более математически логичны, чем полный подсчет<sup>90</sup>. Кроме того, в современных академических

<sup>87</sup> Marusić, A., Bosnjak, L., Jerončić, A. A systematic review of research on the meaning, ethics and practices of authorship across scholarly disciplines [Электронный ресурс]. - 2011. – URL: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0023477>

<sup>88</sup> Frandsen, T. F., & Nicolaisen, J. What is in a name? Credit assignment practices in different disciplines // *Journal of Informetrics*. – 2010. – Vol. 4, No 4. - P. 608-617.

<sup>89</sup> Waltman, L., & Van Eck, N. J. Field-normalized citation impact indicators and the choice of an appropriate counting method // *Journal of Informetrics*. – 2015. – Vol. 9, No 4. - P. 872-894.

<sup>90</sup> Lin, C., Huang, M., Chen, D. The influences of counting methods on university rankings based on paper count and citation count. *Journal of Informetrics*. – 2013. – Vol.7(3), – P.611-621. doi:10.1016/j.joi.2013.03.007

культурах результаты рейтинга часто влияют на национальную политику и распределение ресурсов, более точные методы расчета, такие как прямой подсчет и дробный подсчет, кажутся более подходящими для учреждений, которые придают значение точности и обоснованности оценки методологии<sup>91</sup>.

Более сложные подходы для предоставления кредита авторам на основе их позиции в списке авторов публикации называются методом взвешенного подсчета, который базируется на дробном подсчете с применением различных весовых коэффициентов. Типичная идея данного метода состоит в том, чтобы назначить самый высокий вес первому автору, затем второму, третьему и так далее.

К данной группе относятся:

- гармоничный подсчет<sup>92</sup>;
- арифметический (линейный) подсчет<sup>93</sup>;
- пропорциональный или геометрический подсчет<sup>94</sup>;
- метод подсчета, основанный на золотом числе (метод N.Assimakis и M.Adam)<sup>95</sup>,
- аксиоматический метод подсчета<sup>96</sup>.

Различие распределения авторского кредита по наиболее распространенным подходам указано в таблице 2.2.

---

<sup>91</sup> Gauffriau, M., Larsen, P. O., Maye, I., Roulin-Perriard, A., Von Ins, M. Comparisons of results of publication counting using different methods. *Scientometrics*. – 2008. – Vol.7(1), – P.147-176.

<sup>92</sup> Hagen, N. T. Harmonic allocation of authorship credit: Source-level correction of bibliometric bias assures accurate publication and citation analysis [Электронный ресурс]. - 2008. – URL: doi:10.1371/journal.pone.0004021

<sup>93</sup> Abbas, A. M. Weighted indices for evaluating the quality of research with multiple authorship // *Scientometrics*. – 2011. - Vol. 55, No 1. - P. 107-131.

<sup>94</sup> Egghe, L., Rousseau, R., & Van Hooydonk, G. Methods for accrediting publications to authors or countries: Consequences for evaluation studies // *Journal of the American Society for Information Science*. – 2000. - Vol. 51, No 2. - P. 145-157.

<sup>95</sup> Assimakis, N., Adam, M. (2010). A new author's productivity index: p-index // *Scientometrics*. - 2010. - Vol. 85, No 2. - P. 415-427.

<sup>96</sup> Stallings, J., Vance, E., Yang, J., Vannier, M. W., Liang, J., Pang, L., & Wang, G. Determining scientific impact using a collaboration index / *Proceedings of the National Academy of Sciences of the United States of America*. – 2013. - Vol. 110, No 24. - P. 9680-9685.

Таблица 2.2

**Распределение авторского кредита  
для статей с количеством авторов до 3<sup>97</sup>**

Метод подсчета	Количество авторов	Позиция автора		
		Первый	Второй	Третий
Гармоничный	1	1,0000		
	2	0,6667	0,3333	
	3	0,5455	0,2727	0,1818
Арифметический	1	1,0000		
	2	0,6667	0,3333	
	3	0,5000	0,3333	0,1667
Геометрический	1	1,0000		
	2	0,6667	0,3333	
	3	0,5714	0,2857	0,1429
Дробный	1	1,0000		
	2	0,5000	0,5000	
	3	0,3333	0,3333	0,3333
Метод Асимакиса и Адама	1	1,0000		
	2	0,618	0,382	
	3	0,618	0,236	0,146
Аксиоматический подсчет	1	1,0000		
	2	0,750	0,250	
	3	0,611	0,278	0,111

<sup>97</sup> Hagen, N. T. Harmonic publication and citation counting: sharing authorship credit equitably-not equally, geometrically or arithmetically // *Scientometrics*. – 2010. –Vol. 84, No 3. - P. 785-793.

Гармоничный подсчет одновременно устраняет как инфляционную, так и уравнивающую предвзятость, распределяя количество публикаций и цитирования в соответствии с авторским рангом и количеством соавторов.

При гармоничном подсчете 1-й автор всегда получает вдвое больше кредитов, чем 2-й автор, 2-й автор всегда получает в 1,5 раза больше, чем 3-й, 3-й автор всегда получает в 1,33 раза больше, чем 4-й автор, и так далее<sup>98</sup>.

При гармоничном подсчете первый автор и корреспондент автор воспринимаются как равные участники, и им присуждается равнозначный кредит.

Арифметический подсчет также дает удвоенное количество баллов первому автору, когда соавторов всего два, но не имеет фиксированного коэффициента распределения при увеличении общего количества соавторов<sup>99</sup>.

Геометрический подсчет всегда присваивает первому автору в два раза больше кредита, чем второму, так что первые несколько авторы получают большую часть признания, в то время как остальным достается незначительное признание<sup>100</sup>.

Метод подсчета Асимакиса и Адама основан на так называемом «золотом числе» или «золотой индекс производительности». Он обеспечивает эффективный метод измерения вклада автора. Метод основан на использовании золотого числа  $\phi = 1.6180$  и использует  $1/\phi = 0.6180$  в качестве основы для разделения единицы производительности<sup>101</sup>.

Большинство приведенных методов подсчета не учитывают вклад корреспондент автора, который, как правило, указывается

---

<sup>98</sup> Hagen, N.T. Harmonic publication and citation counting: Sharing authorship credit equitably - not equally, geometrically or arithmetically. *Scientometric*. – 2010. – Vol.84(3), – P.785-793. doi:10.1007/s11192-009-0129-4.

<sup>99</sup> Abbas, A. M. Weighted indices for evaluating the quality of research with multiple authorship. *Scientometrics*. – 2011. – Vol.55(1), – P.107-131.

<sup>100</sup> Egghe, L., Rousseau, R., Van Hooydonk, G. Methods for accrediting publications to authors or countries: Consequences for evaluation studies. *Journal of the American Society for Information Science*. – 2000. – Vol.51(2), – P.145-157.

<sup>101</sup> Assimakis, N., Adam, M. A new author's productivity index: P-index. *Scientometrics*. – 2010. – Vol.85(2), – P. 415-427. doi:10.1007/s11192-010-0255-z

в самом конце списка<sup>102</sup>. Для решения этой проблемы, был предложен метод подсчета, который придает наибольшее значение первому и последнему автору и наименьший вес автору(-ам) в середине списка<sup>103</sup>.

Таким образом, в современной библиометрии ведутся различные дискуссии по выбору наиболее оптимального метода оценки количественных и качественных показателей ученых и организаций для определения их вклада в коллективный труд, а также развитие той или иной научной области. Кроме того, методы оценки производительности ученых и организаций являются действенным инструментом по мониторингу их эффективности и поддержке, в том числе финансовой, наиболее продуктивных. Кроме того, данные методы играют решающую роль в современном управлении и формировании, развитии научной политики государств. Как показал проведенный анализ применение только лишь одних методов анализа количественных и качественных характеристик недостаточно и требуется комплексный подход, что обеспечивает объективный результат.

## **2.2. Техническое описание автоматизированной системы по оценке публикационной активности на основе данных из базы Scopus**

Разработанная авторами система представляет собой самоисполняющееся программное приложение на уровне операционной системы, позволяющее: подключаться к базе данных Scopus по API (Application Programming Interface)-каналу и загружать данные по публикационной активности, цитируемости, индексу Хирша ученых, библиометрическим показателям публикаций; считывать актуальные значения наукометрических показателей

---

<sup>102</sup> Kosmulski, M. The order in the lists of authors in multi-author papers revisited // *Journal of Informetrics*. – 2012. - Vol. 6, No 4. - P. 639-644.

<sup>103</sup> Aziz, N. A., & Rozing, M. P. Profit (p)-index: The degree to which authors profit from co-authors [Электронный ресурс]. - 2013. – URL: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0059814>

авторов и организаций Казахстана на уровне хранимых процедур и скалярных функций реляционной базы данных ранжирования авторов и организаций Казахстана; ранжировать авторов и организации по некоторым показателям публикационной активности, цитируемости, индексу Хирша от большего показателя фактического значения к меньшему и наоборот; подсчитывать общий рейтинговый балл авторов и организаций на основе балльной системы по показателям публикационной активности, цитируемости, индексу Хирша; рейтинговать авторов и организации по общему количеству рейтинговых баллов или по данным показателям балльной системы от большего количества рейтинговых баллов к меньшему.

С помощью системы можно создавать любые статистические отчеты.

● Мониторинг информации по автору – поиск автора, просмотр показателей автора:

- Просмотр аффилиации автора;
- Просмотр общего количества публикаций за определенный период;
- Просмотр количества публикаций в разрезе типа издания (журнал, сборник конференции, глава в монографии, монография);
- Просмотр количества публикаций в разрезе типа документа (журнальная статья, материал конференции, глава в монографии, монография);
- Просмотр количества публикаций в разрезе квартилей, процентилей издания;
- Просмотр количества публикаций без учета публикаций, опубликованных в изданиях, индексация которых прекращена в базе данных Scopus;
- Просмотр количества публикаций единственным автором, корреспондентом автором, первым автором;
- Просмотр количества публикаций с 15 различными аффилиациями;
- Просмотр количества публикаций, профинансированных МОН РК;



- Просмотр соотношения количества журнальных статей к общему количеству публикаций в разрезе квартилей Q1-Q4;
- Просмотр соотношения количества журнальных статей к общему количеству журнальных статей в разрезе квартилей Q1-Q4;
- Просмотр показателя суммарного SJR журналов, в которых опубликованы статьи автора;
- Просмотр показателя суммарного SJR журналов, в которых опубликованы статьи автора, в разрезе квартилей;
- Просмотр общего количества цитирований автора за определенный период, количества цитирований без учета самоцитирований, количества внешних (неаффилированных с организацией) цитирований, количества цитирований без учета аффилированных цитирований, количества цитирований без учета цитирований соавторов, количества цитирований без учета цитирований, пришедших из изданий, индексация которых прекращена в базе данных Scopus, количества цитирований без учета цитирований, пришедших из изданий открытого доступа;
- Просмотр показателя индекса Хирша автора, индекса Хирша без учета самоцитирований, индекса Хирша без учета аффилированных цитирований, индекса Хирша без учета цитирований соавторов, индекса Хирша без учета цитирований, пришедших из изданий, индексация которых прекращена в базе данных Scopus, индекса Хирша без учета цитирований, пришедших из изданий открытого доступа.
- Мониторинг информации по организации – поиск организации, просмотр показателей организации;
- Просмотр количества публикаций авторов организации в разрезе типа издания (журнал, сборник конференции, глава в монографии, монография);
- Просмотр количества публикаций авторов организации в разрезе типа документа (журнальная статья, материал конференции, глава в монографии, монография);
- Просмотр количества публикаций авторов организации в разрезе квартилей, процентилей издания;
- Просмотр количества публикаций авторов организации без учета публикаций, опубликованных в изданиях, индексация которых прекращена в базе данных Scopus;

- Просмотр количества публикаций авторов организации в качестве единственного автора, корреспондент автора, первого автора;
- Просмотр количества публикаций авторов организации с 15 различными аффилиациями;
- Просмотр количества публикаций авторов организации, профинансированных МОН РК;
- Просмотр соотношения количества журнальных статей авторов организации к общему количеству публикаций авторов организации в разрезе квартилей Q1-Q4;
- Просмотр соотношения количества журнальных статей авторов организации к общему количеству журнальных статей авторов организации в разрезе квартилей Q1-Q4;
- Просмотр показателя суммарного SJR журналов, в которых опубликованы статьи авторов организации;
- Просмотр показателя суммарного SJR журналов, в которых опубликованы статьи авторов организации, в разрезе квартилей;
- Просмотр общего количества цитирований авторов организации за определенный период, количества самоцитирований, количества цитирований без учета самоцитирований, количества внешних (неаффилированных с организацией) цитирований, количества цитирований без учета аффилированных цитирований, количества цитирований без учета цитирований соавторов, количества цитирований без учета цитирований, пришедших из изданий, индексация которых прекращена в базе данных Scopus, количества цитирований без учета цитирований, пришедших из изданий открытого доступа;
- Просмотр показателя суммарного индекса Хирша авторов организации, суммарного индекса Хирша без учета самоцитирований, суммарного индекса Хирша без учета цитирований аффилированных с организацией, суммарного индекса Хирша без учета цитирований соавторов, суммарного индекса Хирша без учета цитирований, пришедших из изданий, индексация которых прекращена в базе данных Scopus, суммарного индекса Хирша без учета цитирований, пришедших из изданий открытого доступа;
- Просмотр показателя соотношения количества цитирований авторов организации к количеству уникальных авторов организации;

□ Просмотр показателя соотношения количества публикаций авторов организации к количеству уникальных авторов организации.

Приложение предназначено для загрузки данных из базы данных Scopus и сохранения их в локальной реляционной базе данных вуза. База данных находится под контролем системы управления базами данных Microsoft SQL Server. Сохраненные данные могут использоваться в дальнейшем для построения отчетов и осуществления другой аналитики.

### **Работа с API Scopus**

Для получения данных статей и авторов из базы данных Scopus в качестве первого шага необходимо подключиться к публичному веб-интерфейсу программирования приложений (Web API) Scopus, доступного по адресу <https://api.elsevier.com>. Для осуществления любых запросов к API необходимо в заголовке запроса или параметре метода запроса указать ключ API. Для получения ключа необходимо сделать следующее:

1. Войти на сайт <https://dev.elsevier.com/>;
2. Выбрать в меню «Сгенерировать ключ API».

В результате этих действий будет создан необходимый ключ API в виде строки случайных латинских символов и цифр (например, 1d536b498839e46a8ee3ad57e3b800c4).

Для получения данных об авторах и статьях, приложение обращается к четырем методам, доступным в API:

1. <https://dev.elsevier.com/documentation/AuthorSearchAPI.wadl#d1e33>

Данный метод предназначен для поиска авторов по различным параметрам. В данном приложении используется единственный параметр для поиска – AF-ID (Affiliation ID) – идентификатора организации, к которой относится искомый автор. В качестве значения данного параметра используются Affiliation ID определенной организации, их может быть несколько. В результате запроса можно получить различную информацию по найденным авторам искомой организации. Однако, по данному методу отсутствует информация об индексе Хирша автора. Поэтому единственное предназначение данного метода – получение списка ID авторов, которые будут использованы в качестве параметров для следующего метода.

При отправке запроса в качестве параметра (view) метода можно указать представление (например, STANDARD), которое определяет какие данные будут в ответе на запрос. Возвращаемые данные представлены в следующей таблице:

Поле	Описание	STANDARD
<b>link ref=self</b>	Author Retrieval API URI	X
<b>link ref=scopus-author</b>	Scopus author details URL	X
<b>link ref=scopus-citedby</b>	Scopus Author Cited By Results URL	X
<b>link ref=search</b>	Scopus Search API URL	X
<b>prism:url</b>	Author Retrieval API URI	X
<b>dc:identifier</b>	Author ID	X
<b>eid</b>	Electronic ID	X
<b>orcid</b>	ORCID	X
<b>document-count</b>	Количество документов	X
<b>subject-area</b>	Предметные области	X
preferred-name <b>surname</b>	Предпочитаемая фамилия автора	X
preferred-name <b>given-name</b>	Предпочитаемое имя автора	X
preferred-name <b>initials</b>	Инициалы автора	X
<b>name-variant</b>	Варианты имен автора	X
affiliation-current <b>affiliation-name</b>	Название текущей организации	X
affiliation-current <b>affiliation-city</b>	Город	X
affiliation-current <b>affiliation-country</b>	Страна	X
affiliation-current <b>affiliation-id</b>	ID организации	X
affiliation-current <b>affiliation-url</b>	Affiliation Retrieval API URI	X

2. <https://dev.elsevier.com/documentation/AuthorRetrievalAPI.wadl#d1e322>

Данный метод предназначен для получения подробной информации об авторе по указанному параметру Author ID – уникального идентификатора автора в Scopus. Основные возвращаемые данные: варианты имен автора, список публикаций с указанием электронного идентификатора каждой публикации в базе Scopus (EID), количество цитирующих документов, количество цитирований, название текущей организации автора, город и страна текущей организации автора, индекс Хирша автора, количество соавторов.

Возвращаемые данные представлены в следующей таблице:

Поле	BASIC	METRICS	LIGHT	STANDARD	ENHANCED
<b>dc:identifier</b> ( <i>Author identifier</i> )	X	X	X	X	X
<b>Eid</b> ( <i>Electronic Identifier</i> )	X		X	X	X
<b>Orcid</b> ( <i>ORCID</i> )	X		X	X	X
<b>link ref=scopus-author</b> ( <i>Scopus URL</i> )	X		X	X	X
<b>link ref=self</b> ( <i>Author Retrieval API URI</i> )	X		X	X	X
<b>prism:url</b> ( <i>Author Retrieval API URI</i> )	X	X	X	X	X
<b>link ref=search</b> ( <i>Scopus Search API URL</i> )			X	X	X
<b>document-count</b> ( <i>Количество документов</i> )		X	X	X	X

<b>cited-by-count</b> (Количество цитирующих документов)		X	X	X	X
<b>citations-count</b> (Количество цитирований)		X	X	X	X
preferred-name <b>surname</b> (Предпочитаемая фамилия автора)			X	X	X
preferred-name <b>given-name</b> (Предпочитаемое имя автора)			X	X	X
preferred-name <b>initials</b> (Инициалы автора)			X	X	X
name-variants <b>name-variant</b> (Варианты имен автора)			X	X	X
affiliation-current <b>affiliation-name</b> (Название текущей организации)			X	X	X
affiliation-current <b>affiliation-city</b> (Город текущей организации)			X	X	X
affiliation-current <b>affiliation-country</b> (Страна текущей организации)			X	X	X

affiliation-current <b>affiliation-url</b> ( <i>Affiliation Retrieval API URL</i> )				X	X
affiliation-current <b>affiliation-id</b> ( <i>ID текущей организации</i> )				X	X
affiliation-history <b>affiliation-name</b> ( <i>Историческое название организации</i> )				X	X
affiliation-history <b>affiliation-url</b> ( <i>Affiliation Retrieval API URI</i> )				X	X
affiliation-history <b>affiliation-id</b> ( <i>Исторический ID организации</i> )				X	X
<b>author-profile</b> ( <i>Профиль автора</i> )				X	X
<b>h-index</b> ( <i>Индекс Хирша</i> )		X			X
<b>coauthor-count</b> ( <i>Количество соавторов</i> )		X			X

3. <https://dev.elsevier.com/documentation/ScopusSearchAPI.wadl#d1e33>

Данный метод предназначен для поиска статей (abstracts) по различным параметрам. В данном приложении используется единственный параметр для поиска AU-ID (Author ID) – идентификатора автора, к которому относится искомая статья. В каче-

стве значения данного параметра используются Author ID всех авторов, полученных в результате работы предыдущего метода. В результате запроса можно получить подробную информацию по найденным статьям. Кроме того, данный метод используется для получения информации по статьям, которые ссылаются на найденные до этого статьи (цитирования). Для этого используется параметр REFEID, где указывается EID статьи, для получения информации о цитировании искомой статьи.

Возвращаемые данные представлены в следующей таблице:

Поле	Описание	STANDARD	COMPLETE
<b>link ref=self</b>	Abstract Retrieval API URI	X	X
<b>link ref=scopus</b>	Scopus URL	X	X
<b>link ref=scopus-citedby</b>	Scopus Cited By Results URL	X	X
<b>prism:url</b>	Abstract Retrieval API URI	X	X
<b>dc:identifier</b>	Scopus ID	X	X
<b>eid</b>	Electronic ID	X	X
<b>dc:title</b>	Название	X	X
<b>prism:aggregationType</b>	Тип источника	X	X
<b>subtype</b>	Код типа документа	X	X
<b>subtypeDescription</b>	Описание типа документа	X	X
<b>citedby-count</b>	Количество цитирующих документов	X	X
<b>prism:publicationName</b>	Название источника	X	X
<b>prism:isbn</b>	Идентификатор источника	X	X



<b>prism:issn</b>	Идентификатор источника	X	X
<b>prism:volume</b>	Объем	X	X
<b>prism:issueIdentifier</b>	Выпуск	X	X
<b>prism:pageRange</b>	Страница	X	X
<b>prism:coverDate</b>	Дата публикации	X	X
<b>prism:coverDisplayDate</b>	Дата публикации	X	X
<b>prism:doi</b>	Document Object Identifier	X	X
<b>pii</b>	Publication Item Identifier	X	X
<b>pubmed-id</b>	MEDLINE Identifier	X	X
<b>orcid</b>	ORCID	X	X
<b>dc:creator</b>	Первый автор	X	X
<b>openaccess openaccessFlag</b>	Open Access статус	X	X
<b>affiliation affilname</b>	Название организации	X	X
<b>affiliation affiliation-city</b>	Город организации	X	X
<b>affiliation affiliation-country</b>	Страна организации	X	X
<b>affiliation afid</b>	Affiliation ID		X
<b>affiliation affiliation-url</b>	Affiliation Retrieval API URI		X
<b>affiliation name-variant</b>	Альтернативное название организации		X
<b>author</b>	Полный список авторов		X
<b>dc:description</b>	Описание		X

<b>authkeywords</b>	Ключевые слова		X
<b>article-number</b>	Номер статьи		X
<b>fund-acr</b>	Акроним агентства финансирования		X
<b>fund-no</b>	Идентификатор агентства финансирования		X
<b>fund-sponsor</b>	Название агентства финансирования		X

4. <https://dev.elsevier.com/documentation/AbstractRetrievalAPI.wadl#d1e48>

Данный метод предназначен для получения подробной информации по искомой статье. В качестве параметра используется EID статьи. В результате работы данного метода из ответа на запрос приложение выбирает информацию по корреспондент-авторам, по научным направлениям, а также по финансирующим организациям статьи.

Возвращаемые данные представлены в следующей таблице:

Поле	BASIC	META	META_ABS	REF *	FULL *
<b>eid</b>	X	X	X		X
<b>link ref=scopus</b> (Scopus URL)	X	X	X		X
<b>prism:url</b> (Abstract Retrieval API URI)	X	X	X	X	X
<b>dc:identifier</b> (Scopus ID)	X	X	X	X	X
<b>openaccess openaccessFlag</b> (Open Access статус)	X	X	X		X
<b>link ref=self</b> (Abstract Retrieval API URI)		X	X		X

<b>dc:title</b> (Название)		X	X	X	X
<b>prism:aggregationType</b> (Тип источника)		X	X		X
<b>subtype</b> (Тип документа)		X	X		X
<b>subtypeDescription</b> (Описание типа документа)		X	X		X
<b>citedby-count</b> (Количество цитирований)		X	X	X	X
<b>prism:publicationName</b> (Название источника)		X	X	X	X
<b>prism:isbn</b> (ISBN источника)		X	X		X
<b>prism:issn</b> (ISSN источника)		X	X		X
<b>prism:volume</b> (Объем)		X	X	X	X
<b>prism:issueIdentifier</b> (Идентификатор выпуска)		X	X	X	X
<b>prism:pageRange</b> (Диапазон страниц)		X	X	X	X
<b>prism:coverDate</b> (Дата публикации)		X	X		X
<b>pubmed-id</b> (MEDLINE identifier)		X	X	X	X
<b>prism:doi</b> (Document Object Identifier)		X	X	X	X
<b>article-number</b> (Номер статьи)		X	X	X	X

<b>dc:creator</b> (Первый автор)		X	X		X
<b>affiliation-name</b> (Название организации)		X	X		X
<b>authors</b> (Другие авторы)			X	X	X
<b>afid</b> (Affiliation ID) <b>affiliation-url</b>			X		X
<b>dc:description</b> (Описание)			X		X
<b>intid</b>					X
<b>subject-areas</b> (Предметные области)					X
<b>item</b> (Оригинальный текст)			X		X

Все выгружаемые данные на основе вышеописанных 4-х методов API складываются в локальной реляционной базе данных, таблицы которой формируются автоматически на основе технологии Code First и заполняются через преобразование данных в JSON формате. Такая методика построения и обновления базы данных обеспечивает гибкость при изменении структуры получаемых данных.

Алгоритм работы приложения:

1. Приложение обращается к Author Search API для получения списка авторов организации.
2. Для каждого полученного на предыдущем шаге Author ID происходит обращение к Author Retrieval API и получение информации обо всех авторах.
3. Приложение обращается к Scopus Search API для получения списка статей организации.
4. Для каждой полученной на предыдущем шаге статьи происходит обращение к Scopus Search API и получение информации о статьях, которые ссылаются на данную статью.

5. Для каждой полученной статьи также происходит обращение к Abstract Retrieval API и получение информации о финансирующих организациях, научных направлениях и корреспондент авторах.

6. Сохранение всех полученных данных.

Для работы с базой данных используется библиотека классов ORM (Object Relational Mapper) Entity Framework, которая позволяет работать с таблицами БД, используя уровень абстракции в виде классов данных.

При получении данных с Scopus API, ответ в виде JSON десериализуется в объекты соответствующих классов данных, например объекты класса Author или Abstract. Для работы с JSON используется библиотека Json.NET. (Рис. 2.1).

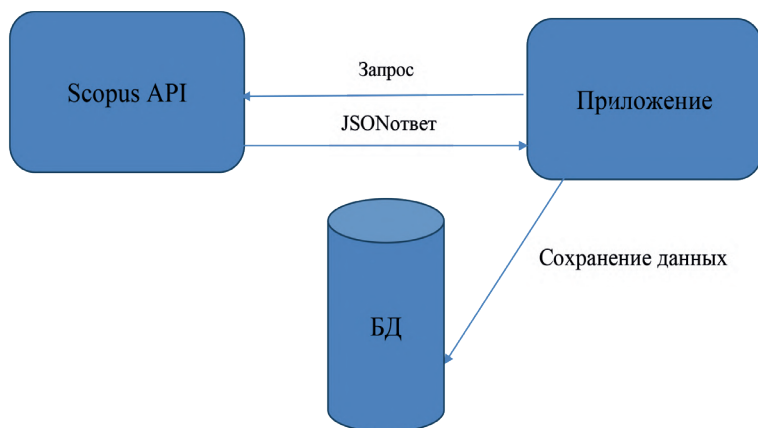


Рисунок 2.1 – Модель работы приложения

Кроме рейтинговой оценки научной производительности ученых и университетов Казахстана по наукометрическим показателям была разработана авторская система ранжирования ученых и организаций на основе авторских показателей. Данные показатели отсутствуют в базе данных Scopus и на платформе SciVal. Так, в авторской системе (наряду со вспомогательными фильтрами по

определению автора на основе ФИО или идентификатора; организации автора; года/диапазона годов публикаций; научного направления; финансируемых МОН РК; тип источника и тип документа) имеется возможность применения следующих основных фильтров:

- Публикации/цитирования в изданиях, индексация которых прекращена в базе данных Scopus;
- Публикации/цитирования не входящих в изданиях, индексация которых прекращена в базе данных Scopus;
- Публикации в изданиях открытого доступа в базе данных Scopus;
- Публикации не входящих в издания открытого доступа в базе данных Scopus;
- Публикации единственным автором в базе данных Scopus;
- Публикации корреспондентом автором в базе данных Scopus;
- Публикации первым автором в базе данных Scopus;
- Публикации с 15 различными аффилиациями в базе данных Scopus;
- Публикации не входящие в список с 15 различными аффилиациями в Scopus.

Введение данных основных фильтров обусловлено тем, что в настоящее время в сфере публикационной активности казахстанских авторов одним из наиболее актуальных вопросов является опубликование статей в так называемых «хищнических» изданиях, которые в короткий период времени исключаются из БД Scopus. В этой связи, для того, чтобы обеспечить объективность рейтинговой оценки научной производительности ученых и университетов Казахстана, был внедрен фильтр по не включению в учет публикаций и цитирований в изданиях, индексация которых прекращена в базе данных Scopus. Кроме того, учитывая тот факт, что большинство изданий, которые исключаются из базы Scopus, являются изданиями открытого доступа, был разработан и введен соответствующий основной фильтр, который позволяет включать/исключать в учет статьи, опубликованные в изданиях открытого доступа.

### 2.3. Структура автоматизированной системы по оценке публикационной активности ученых и организаций Казахстана на основе рейтингового и балльного подхода

Автоматизированная система рейтинговой оценки авторов и организаций (размещенная на сайте <https://scienrating.kaznu.kz/>) на основе данных научной производительности на базе Scopus состоит из четырех модулей:

1. Рейтинг авторов РК;
2. Рейтинг авторов РК – балловая система;
3. Рейтинг организации РК;
4. Рейтинг организации РК – балловая система;

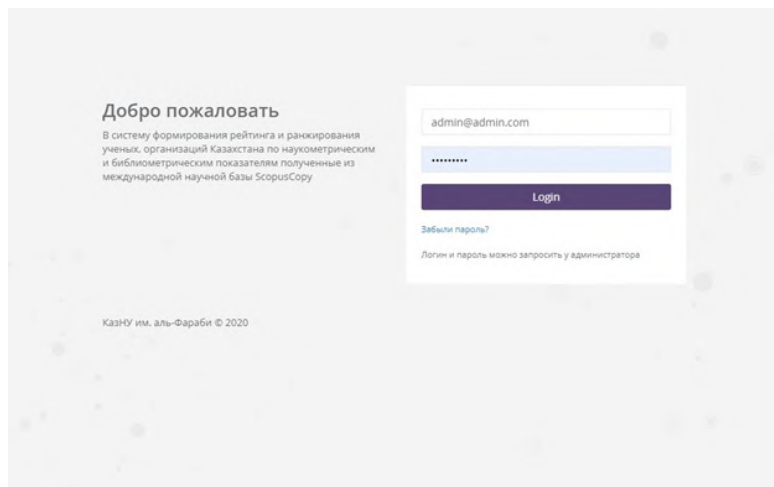


Рисунок 2.2 - Страница авторизации пользователя системы

Данная страница (Рис. 2.2) предназначена для входа в систему, в поля ввода пользователь вводит свои логин и пароль, далее нажимает кнопку для входа и входит в систему, где ему открываются возможности согласно его правам в системе.

Данные в системе формирования рейтинга и ранжирования казахстанских авторов и организации обновляется посредством приложения один раз в месяц.

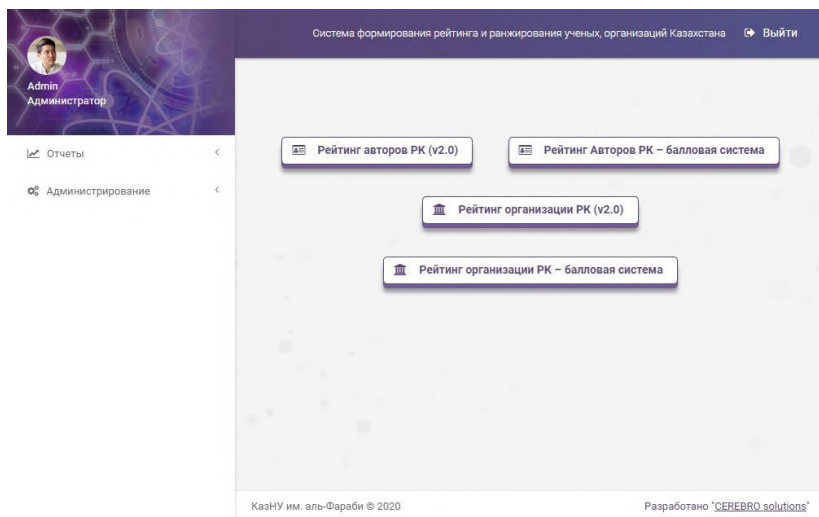


Рисунок 2.3 – Вид главной страницы автоматизированной системы рейтинговой оценки авторов и организаций Казахстана

Главная страница (Рис. 2.3) предназначена для быстрого перехода к основным функционалам сайта, а именно к страницам рейтинга, балловой системы авторов и организаций. Кроме этого, на этой странице имеется левое боковое меню, где также можно использовать ссылки для перехода на страницы «Отчет» и «Администрация».

Таким образом, при наличии прав администратора пользователь получает доступ к страницам администратора, ссылки на эти страницы также отражаются в боковом меню. В правом верхнем углу страницы также находится кнопка для выхода из системы (Рис. 2.3).

### ***2.3.1 Рейтинговая оценка авторов***

Система рейтинговой оценки авторов РК осуществляется путем автоматического расчета фактических значений показателей по публикациям, цитированию, индексу Хирша и библиометрическим показателям изданий, полученных посредством API Scopus.



Сведения в виде количественных показателей отражаются в соответствующих показателях рейтинговой системы (Рис. 2.4). Интерфейс расчета фактических значений наукометрических показателей авторов представлен следующими ролями в системе:

- Панель фильтров;
- Индикаторы;
- Основные и дополнительные параметры;
- Генерация отчета в формате Excel.

ИМЮ автора	Организация автора	Общие цитированные публикации	И <sup>2</sup>	Общие цитированные публикации в статье	Общие цитированные статьи в журнале	Общие цитированные монографии	Общие цитированные статьи в монографии	Общие цитированные монографии	Итого Хард	Суммарный S <sub>h</sub>
Makhovskiy A.	Belarus, Minsk State University, International Centre for Theoretical Physics	274	263	25	0	4	6108	46	108 109	
Kuznetsov A.	Institute of Nuclear Physics, National Nuclear Center of the Republic of Kazakhstan	227	230	21	0	1	2058	38	127 261	
Makarov Z.	Al-Farabi Kazakh National University	247	166	11	0	0	1053	14	64 317	
Levashenko T.	Al-Farabi Kazakh National University	233	175	16	0	1	2465	20	102 283	
Naibinov A.	Kazakh-British Technical University, Department of Mathematics and Statistics	206	102	21	0	1	1109	22	141 616	
Gavrilov V.	Al-Farabi Kazakh National University, Department of Theoretical and Nuclear Physics	178	107	20	0	2	2059	20	106 843	
Chernyshov V.	Al-Farabi Kazakh National University	164	138	10	0	3	866	17	117 772	
Kudabergina S.	Belarusian University	147	116	17	1	4	2809	21	63 605	
Mak G.	Al-Farabi Kazakh National University	136	126	6	0	1	1607	26	83 469	
Shalimov R.	Institute of Nuclear Physics, National Nuclear Center of the Republic of Kazakhstan	120	96	20	0	0	403	16	41 073	

Рисунок 2.4. – Интерфейс автоматизированной системы рейтинга авторов РК

**Панель фильтров.** В левом меню расположен набор различных фильтров для поиска необходимых пользователю данных (Рис. 2.5):

- ФИО – фамилия, имя, отчество автора в латинской транслитерации, согласно базе данных Scopus. Возможность выбора до 100 авторов по данному фильтру;
- Scopus Author ID – идентификатор автора в базе данных Scopus. Возможность выбора до 100 авторов по данному фильтру;
- Период публикации – диапазон годов публикации научных статей и других типов публикаций выбранного автора(-ов);
- Организация – аффилиация, указанная автором(-ами) искомыми публикаций;

- Тип источника – выбор типа издания, в котором опубликована публикация(-ии) искомого автора(-ов) (Journal, Conference Proceeding, Book, Book Series, Trade Journal);
- Тип документа – поиск публикаций искомого автора(-ов) по типам документов «Article», «Review Article», «Conference Paper», «Book», «Book Chapter» по базе данных Scopus;
- Научное направление – предметная область искомых публикаций;
- Финансирующая организация МОН РК – признак наличия у публикации среди спонсоров Министерства образования и науки Республики Казахстан.

**Фильтр**

**ФИО**  
ФИО

**Scopus Author ID**  
ID Author

**Период публикации**  
1 990 год 2 021 год

**Организация**  
Выберите организации

**Тип источника**  
Выбрать

**Тип документа**  
Выбрать

**Научное направление**  
Выбрать

**Финансирующая организация МОН РК**  
Выбрать

**Отфильтровать**

Рисунок 2.5 – Вид окна панели фильтров рейтинговой системы авторов РК

**Панель индикаторов** – это дополнительные фильтры, расположенные в верхней части автоматизированной системы. Состоит из следующих индикаторов (Рис. 2.6):

- Публикации/цитирования в изданиях, индексация которых прекращена в базе данных Scopus;
- **Публикации/цитирования, не входящих в издания, индексация которых прекращена в базе данных Scopus;**
- Публикации в изданиях открытого доступа;
- Публикации, не входящих в издания открытого доступа;
- Публикации единственным автором;
- Публикации корреспондентом автором;
- Публикации первым автором;
- Публикации с 15 различными аффилиациями;
- Публикации, не входящие в список с 15 различными аффилиациями.

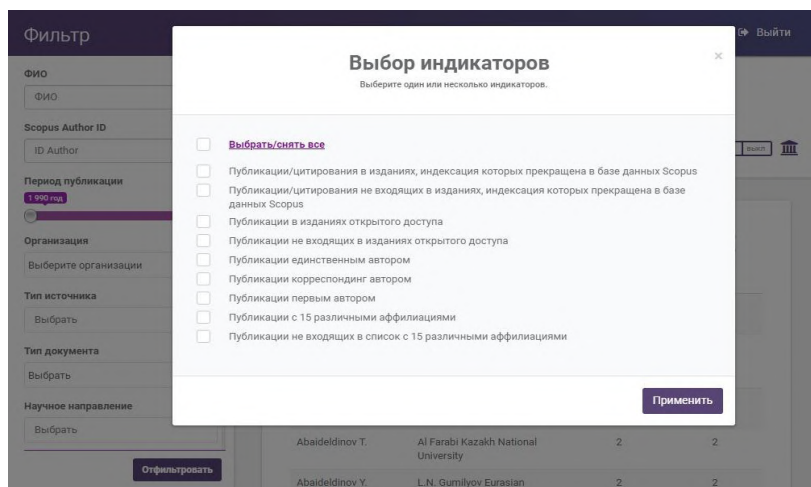


Рисунок 2.6 – Вид окна панели индикаторов рейтинговой системы авторов РК

### ***Основные и дополнительные параметры***

К основным параметрам относятся ряд параметров, дающих значения наиболее важных наукометрических показателей авторов по данным Scopus:

- ФИО автора (указывается ФИО одного или нескольких авторов, выбранных с помощью Scopus Author ID или ФИО);
- Организация автора (согласно профиля организации по AffiliationID в базе данных Scopus);
- Общее количество публикаций (автоматизированная система суммирует данные следующих параметров типов документов, полученных из базы данных Scopus: общее количество журнальных статей, статей в материалах конференции, монографий и глав монографий);
- Общее количество журнальных статей (автоматизированная система определяет общее количество публикаций по типу документа «Article» и «Review» по базе данных Scopus, в изданиях, которые имеют тип «Journal»);
- Общее количество статей в материалах конференций (автоматизированная система определяет общее количество публикаций по типу документа Conference Paper в базе данных Scopus);
- Общее количество монографий (автоматизированная система определяет общее количество публикаций по типу документа Book в базе данных Scopus);
- Общее количество глав в монографиях (автоматизированная система определяет общее количество публикаций по типу документа Book Chapter в базе данных Scopus);
- Общее количество цитирований (автоматизированная система определяет общее количество ссылок, поступивших на публикацию(-ии) выбранного автора(-ов) из публикаций других авторов в базе данных Scopus);
- Индекс Хирша (высчитывается по формуле  $n$  количество публикаций автора процитированных не менее  $n$  количество раз каждая в базе Scopus);
- Суммарный SJR (суммируется количество SJR журналов, в которых опубликованы статьи выбранного автора(-ов)).

ФИО автора	Организация автора	Общее количество публикаций	Общее количество журнальных статей	Общее количество статей в материалах	Общее количество монографий	Общее количество глав в монографиях	Общее количество цитирований	Индекс Хирша	Суммарный SJR
Myrzakulov R.	Ratbay Myrzakulov Eurasian International Centre for Theoretical Physics	274	243	25	0	4	6159	45	188.109
Kozlovskiy A.	Institute of Nuclear Physics, National Nuclear Center of the Republic of Kazakhstan	257	230	21	0	1	2960	30	137.781

Рисунок 2.7 – Вид окна основных параметров рейтинговой системы авторов РК

Рядом с каждым названием показателей основных параметров находятся кнопки (вверх/вниз), которые позволяют ранжировать количественный показатель выбранного индикатора от максимального значения к минимальному, или от минимального значения к максимальному (Рис. 2.7).

Квартиль – метрика, применяемая в базе данных Scopus для ранжирования представленных в определенном научном направлении изданий на 4 категории, в зависимости от показателя процентиля издания, согласно рассматриваемого научного направления. Квартиль 1 (или Q1) – диапазон показателя процентиля от 75 до 99, квартиль 2 (Q2) – процентиль от 50 до 74, квартиль 3 (Q3) – процентиль от 25 до 49, квартиль 4 (Q4) – процентиль от 1 до 24.

**Дополнительные параметры** позволяют проводить детальный анализ наукометрических показателей автора(-ов). Поэтому справа от панели индикаторов будет кнопка «Дополнительные параметры», и она реализована с помощью функции MultiSelect. Дополнительные параметры состоят из следующих показателей:

- **Анализ публикаций в целом в разрезе квартилей Q1-Q4:**
  - Количество публикаций в квартиле 1 (диапазон процентиля: 75-99);
  - Количество публикаций в квартиле 2 (диапазон процентиля: 50-74);
  - Количество публикаций в квартиле 3 (диапазон процентиля: 25-49);

– Количество публикаций в квинтиле 4 (диапазон процентиля: 1-24);

• **Анализ журнальных статей в разрезе квинтилей Q1-Q4:**

– Количество статей в квинтиле 1 (диапазон процентиля: 75-99);

– Количество статей в квинтиле 2 (диапазон процентиля: 50-74);

– Количество статей в квинтиле 3 (диапазон процентиля: 25-49);

– Количество статей в квинтиле 4 (диапазон процентиля: 1-24);

• **Соотношение количества журнальных статей к общему количеству публикаций в разрезе квинтилей Q1-Q4:**

– Соотношение количества журнальных статей в квинтиле 1 к общему количеству публикаций;

– Соотношение количества журнальных статей в квинтиле 2 к общему количеству публикаций;

– Соотношение количества журнальных статей в квинтиле 3 к общему количеству публикаций;

– Соотношение количества журнальных статей в квинтиле 4 к общему количеству публикаций;

• **Соотношение количества журнальных статей к общему количеству журнальных статей в разрезе квинтилей Q1-Q4:**

– Соотношение количества журнальных статей в квинтиле 1 к общему количеству журнальных статей в квинтилях 1-4;

– Соотношение количества журнальных статей в квинтиле 2 к общему количеству журнальных статей в квинтилях 1-4;

– Соотношение количества журнальных статей в квинтиле 3 к общему количеству журнальных статей в квинтилях 1-4;

– Соотношение количества журнальных статей в квинтиле 4 к общему количеству журнальных статей в квинтилях 1-4;

В базе данных Scopus основным показателем качества журналов является SCImago Journal Rank (SJR)<sup>104</sup>. Важность внедрения одним из основных показателей оценки эффективности публикационной активности ученых и организаций на основе данных из базы Scopus суммарного SJR связана со следующим. В журналах с

---

<sup>104</sup> Kotsemir, M., & Shashnov, S. (2017). Measuring, analysis and visualization of research capacity of university at the level of departments and staff members. *Scientometrics*, 112(3), 1659-1689. doi:10.1007/s11192-017-2450-7

сомнительной репутацией, где наблюдается огромное количество самоцитирований, и цитирований, пришедших из журналов, с которыми и/или с издательствами которых (группа журналов) заключен негласный договор о перекрестном цитировании, показатель SJR остается не высоким. Так как индекс SJR учитывает авторитет и позицию журнала в определенной научной области, из которой пришло цитирование на статью рассматриваемого журнала. Примерно аналогичные расчеты внедрены в авторской автоматизированной системе, где суммарный SJR умножается на 1000 баллов. В итоге, чем выше SJR, тем больше баллов получит автор.

● **Суммарный SJR:**

- Суммарный SJR публикаций в Q1;
- Суммарный SJR публикаций в Q2;
- Суммарный SJR публикаций в Q3;
- Суммарный SJR публикаций в Q4;

● **Количество журнальных статей в зависимости от процентиля:**

- Количество публикаций в журналах с процентилем выше 25;
- Количество публикаций в журналах с процентилем выше 35;
- Количество публикаций в журналах с процентилем выше 45;
- Количество публикаций в журналах с процентилем выше 50;
- Количество публикаций в журналах с процентилем выше 97;

● **Анализ по цитированию:**

- Общее количество самоцитирований;
- Общее количество без самоцитирований;
- Количество внешних (неаффилированных с организацией) цитирований;

- Количество цитирований без учета цитирований всех соавторов;

● **Соотношение цитирований к количеству авторов публикации.**

Индекс Хирша – количественная мера производительности труда ученого, организации или страны в целом, основанная на количестве публикаций и количестве цитирований этих публикаций. Вычисляется по формуле  $n$  количество публикаций автора/организации/страны, процитированных не менее  $n$  количество раз каждая в базе Scopus.

Индекс Хирша имеет некоторые ценные преимущества<sup>105</sup>, можно также подчеркнуть некоторые явные недостатки, такие как тот факт, что H-индекс изменяется в идеальном соответствии с научной базой данных, использованной для его расчета<sup>106</sup>, он не отражает реальный вклад ученого в публикацию (т. е. положение ученого в списке авторов не учитывается)<sup>107</sup>.

● **Индекс Хирша:**

- Индекс Хирша без учета самоцитирований;
- Индекс Хирша без учета цитирований всех соавторов;
- Индекс Хирша без учета цитирований в журналах открытого доступа;
- Индекс Хирша без учета цитирований в исключенных из базы данных журналах;
- Индекс Хирша без учета самоцитирований, цитирований всех соавторов, цитирований аффилированных с организацией, цитирований в журналах открытого доступа и исключенных из базы данных журналах (Рис. 2.8).

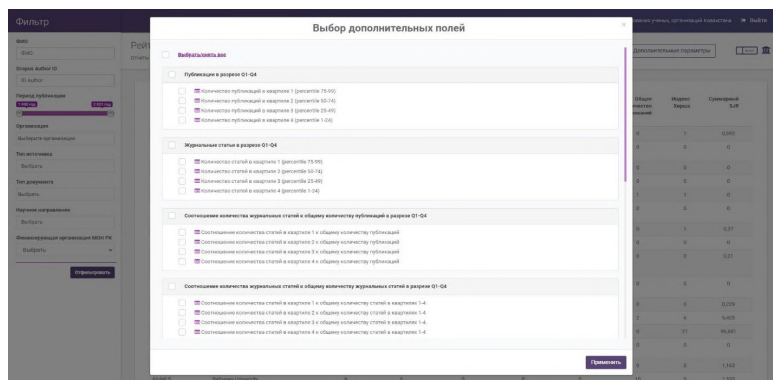


Рисунок 2.8 – Вид окна дополнительных параметров рейтинговой системы авторов ПК

<sup>105</sup> Lippi G, Borghi L. A short story on how the H-index may change the fate of scientists and scientific publishing. *Clin Chem Lab Med* 2014;52:e1-3.

<sup>106</sup> Lippi G, Mattiuzzi C. The challenges of evaluating scientists by H-index and citations in different biomedical research platforms. *Clin Chim Acta* 2013;421:57

<sup>107</sup> Lippi, G., & Mattiuzzi, C. (2017). Scientist impact factor (SIF): A new metric for improving scientists' evaluation? *Annals of Translational Medicine*, 5(15) doi:10.21037/atm.2017.06.24



**Генерация отчета в формате Excel** предназначено для загрузки наукометрических показателей искомого автора(-ов) по выбранным фмильграм, параметрам, дополнительным параметрам в формате XLS/CSV (Рис. 2.9).

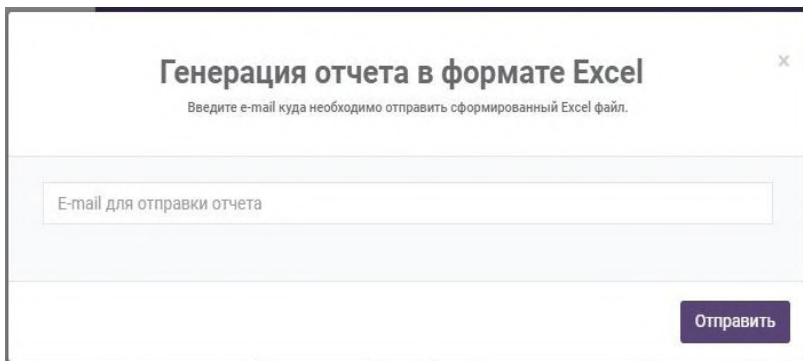


Рисунок 2.9 – Вид окна генерация отчета в формате Excel рейтинговой системы авторов РК

В правом верхнем углу системы расположен слайдер с функцией *вкл/выкл*, предназначенный для отображения публикаций, опубликованных автором с текущей аффилиацией. Необходимость внедрения данной функции связано с тем, что в публикационной истории автора могут быть публикации, в которых указаны различные аффилиации, для выбора публикаций с текущей аффилиацией автора, указанной в получаемых через приложение данных из базы Scopus в качестве параметра «Current Affiliation». При нажатии на слайдер в таблице отображается список публикаций выбранного автора(-ов) с текущей его аффилиацией.

Также, внизу страницы со списком авторов имеется возможность изменить количество одновременно отображаемых найденных записей, там же есть возможность переключения страниц среди найденных результатов (Рис. 2.10).

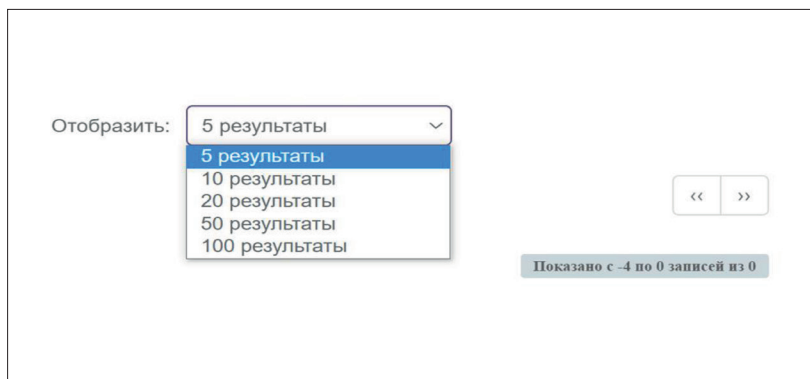


Рисунок 2.10 – Вид нижней части рейтинговой системы авторов РК

### ***2.3.2 Рейтинг Авторов – балловая система.***

Модуль «Рейтинг авторов РК – балловая система» предназначен для просмотра результатов авторов в виде численных показателей (фактических значений), умноженных на соответствующий рейтинговый балл показателя (Рис. 2.11).

В рейтинге автора разработанный нами автоматизированной системе исключен учет публикаций (соответственно пришедшие на данные публикации цитирования), в которых авторы представлены 15 и более аффилиациями (аналогично методологии рейтингового агентства QS), так как данные статьи искажают статистику автора/организации, в связи с тем, что подобного рода статьи носят мультидисциплинарный характер с участием большого количества авторов и организаций; а, во-вторых, большое количество авторов не показывает и затрудняет определение вклада каждого автора.

Функционал страницы аналогичен странице рейтинга авторов.

**Рейтинг Авторов РК – балловая система**

Отчеты / Рейтинг Авторов РК – балловая система

ФИО автора	Организация автора	Общее количество статей в материалах	Общее количество монографий	Общее количество глав в монографиях	Количество публикаций финансируемых МОН РК	Количество статей в журнале "1" (percentile)
Usenov Y.	Institute of Applied Sciences and Information Technologies	50	0	0	40	400

Отобразить: 5 результатов

Вывести (с 1 по 5 страниц из 8)

---

**Рейтинг Авторов РК – балловая система**

Отчеты / Рейтинг Авторов РК – балловая система

Количество цитирований без учета цитирований	Индекс Хирша без учета самоцитирований	Индекс Хирша без учета цитирований	Индекс Хирша без учета цитирований с организацией	Индекс Хирша без учета цитирований	Индекс Хирша без учета цитирований	Индекс Хирша без учета самоцитирований, цитирований влск	Итоговый рейтинговый балл автора
370	60	60	90	180	240	150	6981 000

Рисунок 2.11 – Вид страницы балловой системы рейтинга авторов РК

### 2.3.3 Рейтинговая оценка организаций

Автоматизированная система рейтинга организаций РК (Рис. 2.12) предназначена для просмотра и анализа данных, полученных посредством приложения, по различным показателям организаций Казахстана в базе данных Scopus.

Интерфейс расчета фактических значений наукометрических показателей организаций представлен следующими модулями в системе:

- Панель фильтров;
- Индикаторы;
- Основные и дополнительные параметры;
- Генерация отчета в формате Excel.

Система формирования рейтинга и ранжирования ученых, организаций Казахстана

Выйти

Фильтр

Период публикации: 1980-2020

ВУЗы: Не выбран

Тип источника: Выбрать

Тип документа: Выбрать

Научное направление: Выбрать

Финансирующая организация МОН РК: Выбрать

Оформить

### Рейтинг Организаций

Отчеты / Рейтинг Организаций

Блок Индикаторы Дополнительные параметры

Название организации	Общее количество публикации	Общее количество журнальных статей	Общее количество статей в материалах	Общее количество монографии	Общее количество глав в монографиях	Общее количество цитирований	Суммарный Индекс Хирша всех авторов	Суммарный
Al-Farabi Kazakh National University	6702	5007	1099	99	87	19319	3874	742
L.N. Gumilyov Eurasian National University	3508	2697	563	17	16	9664	1734	214
Satbayev University	1725	1087	493	29	24	3841	1092	128
Institute of Nuclear Physics, National Nuclear Center of the Republic of Kazakhstan	679	500	66	0	0	2952	524	87
Institute of Combustion Problems	1	1	0	0	0	0	12	0.1

Отображать: 5 результатов

1

Рисунок 2.12 – Интерфейс автоматизированной системы рейтинга организаций РК

**Панель фильтров.** В левом меню расположен набор различных фильтров для поиска необходимых пользователю данных (Рис. 2.13):

- Период публикации – диапазон годов публикации научных статей и других типов публикаций выбранной организации(-ий);
- ВУЗы – список названий вузов, указанных на английском языке, согласно базе данных Scopus, с возможностью выбора нескольких организаций через функцию *MultiSelect*;
- Тип источника – выбор типа издания, в котором опубликована публикация(-ии) искомой организации(-ий) (Journal, Conference Proceeding, Book, Book Series, Trade Journal);
- Тип документа – поиск публикаций искомой организации(-ий) по типам документов «Article», «Review Article», «Conference Paper», «Book», «Book Chapter» по базе данных авторов Scopus;
- Научное направление – предметная область искомых публикаций;
- Финансирующая организация МОН РК – признак наличия у публикации среди спонсоров Министерства образования и науки Республики Казахстан.

The image shows a web-based filter panel titled "Фильтр" (Filter). It contains several sections for filtering data:

- Период публикации** (Publication period): A range slider set from "1 990 год" to "2 021 год".
- ВУЗы** (Universities): A dropdown menu with the text "Не выбран" (Not selected).
- Тип источника** (Source type): A dropdown menu with the text "Выбрать" (Select).
- Тип документа** (Document type): A dropdown menu with the text "Выбрать" (Select).
- Научное направление** (Scientific direction): A dropdown menu with the text "Выбрать" (Select).
- Финансирующая организация МОН РК** (Funding organization of the Ministry of Education of the Republic of Kazakhstan): A dropdown menu with the text "Выбрать" (Select).

At the bottom of the panel is a purple button labeled "Отфильтровать" (Filter).

Рисунок 2.13 – Вид окна панели фильтров рейтинговой системы организаций РК

**Панель индикаторов** – это дополнительные фильтры, расположенные в верхней части автоматизированной системы. Состоит из следующих индикаторов (Рис. 2.14):

- Публикации/цитирования в изданиях, индексация которых прекращена в базе данных Scopus;
- *Публикации/цитирования, не входящих в издания, индексация которых прекращена в базе данных Scopus;*
- Публикации в изданиях открытого доступа;
- Публикации, не входящих в издания открытого доступа;
- Публикации единственным автором;
- Публикации корреспондентом автором;
- Публикации первым автором;
- Публикации с 15 различными аффилиациями;
- Публикации, не входящие в список с 15 различными аффилиациями.

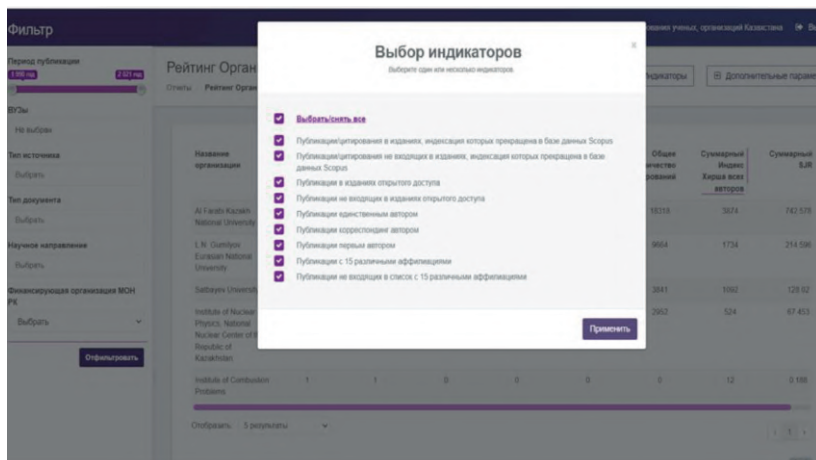


Рисунок 2.14 – Вид окна панели индикаторов рейтинговой системы организаций РК

### ***Основные и дополнительные параметры***

Рейтинговая система организаций состоит из основных 8 показателей:

- Название организации (название организации на английском языке согласно Affiliation ID в базе данных Scopus);
- Общее количество публикаций (автоматизированная система суммирует данные следующих параметров типов документов, полученных из базы данных Scopus: общее количество журнальных статей, статей в материалах конференции, монографий и глав монографий);
- Общее количество журнальных статей (автоматизированная система определяет общее количество публикаций по типу документа «Article» и «Review» по базе данных Scopus, в изданиях, которые имеют тип «Journal»);
- Общее количество статей в материалах конференций (автоматизированная система определяет общее количество публикаций по типу документа Conference Paper в базе данных Scopus);

- Общее количество монографий (автоматизированная система определяет общее количество публикаций по типу документа Book в базе данных Scopus);
- Общее количество глав в монографиях (автоматизированная система определяет общее количество публикаций по типу документа Book Chapter в базе данных Scopus);
- Общее количество цитирований (автоматизированная система определяет общее количество ссылок, поступивших на публикации авторов выбранной организации из публикаций других авторов в базе данных Scopus);
- Суммарный Индекс Хирша всех авторов организации (суммируется количество показателей Индекса Хирша всех авторов выбранной организации в базе Scopus);
- Суммарный SJR (суммируется количество SJR журналов, в которых опубликованы статьи всех авторов выбранной организации) (Рис. 2.15).

Название организации	Общее количество публикаций	Общее количество журнальных статей	Общее количество статей в материалах	Общее количество монографий	Общее количество глав в монографиях	Общее количество цитирований	Суммарный Индекс Хирша всех авторов организации	Суммарный SJR
Al Farabi Kazakh National University	6762	5007	1099	99	87	18318	3874	742.578
L.N. Gumilyov Eurasian National University	3508	2697	563	17	16	9694	1734	214.596
Satbayev University	1725	1087	493	29	24	3841	1092	128.02
Institute of Nuclear Physics, National Nuclear Center of the Republic of Kazakhstan	679	550	68	0	0	2952	524	87.453
Institute of Combustion Problems	1	1	0	0	0	0	12	0.188

Рисунок 2.15 – Вид окна панели индикаторов рейтинговой системы организаций РК

Рядом с каждым названием показателей основных параметров находятся кнопки (вверх/вниз), которые позволяют ранжировать количественный показатель выбранного индикатора от максимального значения к минимальному, или от минимального значения к максимальному.

**Дополнительные параметры** позволяют проводить детальный анализ наукометрических показателей выбранной организации. Поэтому справа от панели индикаторов будет кнопка «Дополнительные параметры», и она реализована с помощью кнопок

ки функции MultiSelect. Дополнительные параметры состоят из следующих показателей:

● **Анализ публикаций в целом в разрезе квартилей Q1-Q4:**

– Количество публикаций в квартиле 1 (диапазон процентиля: 75-99);

– Количество публикаций в квартиле 2 (диапазон процентиля: 50-74);

– Количество публикаций в квартиле 3 (диапазон процентиля: 25-49);

– Количество публикаций в квартиле 4 (диапазон процентиля: 1-24);

● **Анализ журнальных статей в разрезе квартилей Q1-Q4:**

– Количество статей в квартиле 1 (диапазон процентиля: 75-99);

– Количество статей в квартиле 2 (диапазон процентиля: 50-74);

– Количество статей в квартиле 3 (диапазон процентиля: 25-49);

– Количество статей в квартиле 4 (диапазон процентиля: 1-24);

● **Соотношение количества журнальных статей к общему количеству публикаций в разрезе квартилей Q1-Q4:**

– Соотношение количества журнальных статей в квартиле 1 к общему количеству публикаций;

– Соотношение количества журнальных статей в квартиле 2 к общему количеству публикаций;

– Соотношение количества журнальных статей в квартиле 3 к общему количеству публикаций;

– Соотношение количества журнальных статей в квартиле 4 к общему количеству публикаций;

● **Соотношение количества журнальных статей к общему количеству журнальных статей в разрезе квартилей Q1-Q4:**

– Соотношение количества журнальных статей в квартиле 1 к общему количеству журнальных статей в квартилях 1-4;

– Соотношение количества журнальных статей в квартиле 2 к общему количеству журнальных статей в квартилях 1-4;

– Соотношение количества журнальных статей в квартиле 3 к общему количеству журнальных статей в квартилях 1-4;

– Соотношение количества журнальных статей в квартиле 4 к общему количеству журнальных статей в квартилях 1-4;



- **Суммарный SJR:**

- Суммарный SJR публикаций в Q1;
- Суммарный SJR публикаций в Q2;
- Суммарный SJR публикаций в Q3;
- Суммарный SJR публикаций в Q4;

- **Количество журнальных статей в зависимости от процентиля:**

- Количество публикаций в журналах с процентилем выше 25;
- Количество публикаций в журналах с процентилем выше 35;
- Количество публикаций в журналах с процентилем выше 45;
- Количество публикаций в журналах с процентилем выше 50;
- Количество публикаций в журналах с процентилем выше 97;

- **Анализ по цитированию:**

- Общее количество самоцитирований;
- Общее количество без самоцитирований;
- Количество внешних (неаффилированных с организацией) цитирований;

- Количество цитирований без учета цитирований всех соавторов;

- **Соотношение цитирований к количеству авторов:**

- Соотношение количества внешних (неаффилированных с организацией) цитирований к общему количеству цитирований публикаций выбранной организации(-ий);

- Соотношение количества цитирований без учета цитирований всех соавторов к общему количеству цитирований публикаций выбранной организации(-ий);

- Соотношение количества цитирований без самоцитирований и цитирований всех соавторов к общему количеству цитирований публикаций выбранной организации(-ий);

- **Суммарный Индекс Хирша:**

- Суммарный Индекс Хирша;
- Суммарный Индекс Хирша без учета самоцитирований авторов организации;

- Суммарный Индекс Хирша без учета цитирований всех соавторов организации;

- Суммарный Индекс Хирша без учета цитирований аффилированных с организацией;

– Суммарный Индекс Хирша без учета самоцитирований, цитирований всех соавторов организации, цитирований аффилированных с организацией, цитирований в журналах открытого доступа, цитирований в исключенных из базы данных журналах;

● **Количество уникальных авторов организации.**

● **Соотношение количества публикаций/цитирований к количеству уникальных авторов:**

– Соотношение количества цитирований к количеству уникальных авторов выбранной организации;

– Соотношение количества публикаций к количеству уникальных авторов выбранной организации;

● **Доля авторов в числе соавторов из одной организации** (Рис. 2.16).

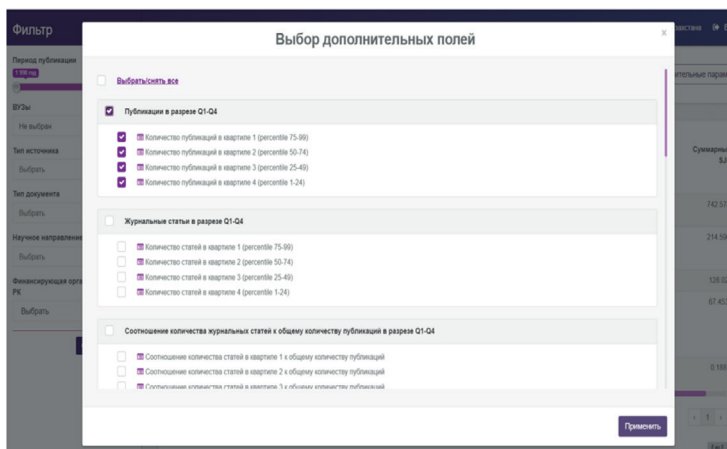


Рисунок 2.16 – Вид окна дополнительных параметров рейтинговой системы организаций РК

**Генерация отчета в формате Excel** предназначено для загрузки наукометрических показателей искомой организации(-ий) по выбранным фмилтрам, параметрам, дополнительным параметрам в формате XLS/CSV (Рис. 2.17).

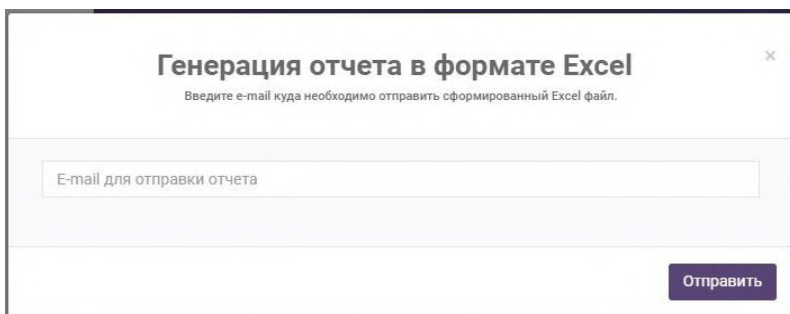


Рисунок 2.17 – Вид окна генерация отчета в формате Excel рейтинговой системы организаций РК

### 2.3.4 Рейтинг Организаций – балловая система.

Модуль «Рейтинг Организаций РК – балловая система» предназначен для просмотра результатов организаций в виде численных показателей (фактических значений), умноженных на соответствующий рейтинговый балл показателя (Рис. 2.18).

Название организации	количество статей в авторских конференциях	Общее количество монографий	Общее количество глав в монографиях	Суммарный SJR	Количество публикаций финансируемых МОН РК	Итоговый рейтинговый балл организации
Al Farabi Kazakh National University	54950	9900	4350	74257.800	9450	152917.800
L.N. Gumilyev Eurasian National University	29150	1700	800	21459.600	3680	55789.600
Sabayev University	24850	2900	1200	12802.00	1900	43432.00
Institute of Nuclear Physics, National Nuclear Center of the Republic of Kazakhstan	3400	0	0	6745.300	960	11105.300
Institute of Combustion Problems	0	0	0	18.800	0	18.800

Рисунок 2.18 – Вид страницы балловой системы рейтинга организаций РК

Таким образом, разработанная авторами автоматизированная система позволяет анализировать количественные и качественные показатели публикационной активности казахстанских уче-

ных и организаций, а также на основе рейтинговой и балловой систем формировать рейтинг авторов и организаций по результатам фактической производительности или набранных баллов.

#### **2.4. Список статей, опубликованных в рамках исследования:**

Kudaibergenova R., Uzakbay S., Ramadinkyzy K., Makanova A., Kistaubayev E., Dussekeev R., Smagulov K. Managing Publication Change at Al-Farabi Kazakh National University: A Case Study // *Scientometrics*. – 2021. - <https://doi.org/10.1007/s11192-021-04139-y> (Scopus, процентиль по CiteScore – 95)

Рамадинкызы К., Ұзақбай С.А. Применение информационно-аналитических платформ для оценки научно-исследовательской деятельности сотрудников университета // *3i: intellect, idea, innovation – интеллект, идея, инновация*. – 2020. – № 3. – С.104-113. – Импакт-фактор журнала по КазБЦ – 0,083. (КОКСОН МОН РК)

Сеилханова М.Е., Омиралиева Г.К. Жоғары оқу орындарын зерттеу университеттерінде айналдырудың өзекті мәселелері // *Вестник КазНПУ им. Абая. Серия «Социологические и политические науки»*. – 2020. – № 3(71). – С.142-150. – Импакт-фактор журнала по КазБЦ – 0,094. (КОКСОН МОН РК)

Слямова Э.С. Жоғары оқу орындарындағы зияткерлік капиталды бағалаудың теңгерімді көрсеткіштер жүйесі // *Вестник КазНПУ им. Абая. Серия «Социологические и политические науки»*. – 2020. – № 4(72). – С.145-150. – Импакт-фактор журнала по КазБЦ – 0,094. (КОКСОН МОН РК)

Маканова А.У., Тургинбаева А.Н., Смагулова Г.С., Сейтхамзина Г.Ж. Государственное регулирование научно-исследовательской деятельности: зарубежный опыт // *Статистика, учет и аудит*. – 2020. – №4(79). – С.175-180. Импакт-фактор журнала по КазБЦ – 0,03. (КОКСОН МОН РК)

Маканова А.У., Тургинбаева А.Н., Смагулов К.Е. Повышение привлекательности научно-исследовательской деятельности в Казахстане: текущее состояние и предпринимаемые меры // *Экономика и статистика*. – 2020. – №4. – С. 96-105. (КОКСОН МОН РК)

Маканова А.У., Тургинбаева А.Н., Смагулова Г.С., Смагулов К.Е. Управленческие аспекты трансформации исследовательских университетов: зарубежный опыт // Вестник КазУЭФМТ. – 2020. – № 4(41). – С.182-190. Импакт-фактор журнала по КазБЦ – 0,171. (КОКСОН МОН РК)

Кудайбергенова Р.Е., Смагулов К.Е., Омиралиева Г.К. Казахстанские ВУЗы в мировых рейтингах: тренды, движущие силы и проблемы // Вестник КазНУ. Серия педагогические науки. – 2021. – № 3(68). – С.94-113. Импакт-фактор журнала по КазБЦ – 0,09. (КОКСОН МОН РК)

Омиралиева Г.К., Сеилханова М.Е. Қазақстанның ғылым және зияткерлік капитал саласындағы статистикалық деректерді талдау // Глобальная наука и инновация 2020: Центральная Азия. Спец выпуск по итогам X Международной научно-практической конференции «Global science and innovations 2020. Central Asia». – 2020. – № 5(10). – С. 40-45.

Маканова А.У. Современные вызовы опубликования научных статей казахстанских авторов в международных научных журналах / Сб. материалов VI Международной научно-практической конференции «Наука и образование в современном мире: вызовы XXI века» / Сост.: Е. Ешім, Е. Абиєв. – Нур-Султан, 2020. – С.60-65.

Маканова А.У., Смагулов Г.С., Тургинбаева А.Н. Механизмы оценки качества научных исследований: опыт зарубежных стран и Казахстана / Сб. материалов межд. науч. конф. «Модели и методы повышения эффективности инновационных исследований» / Под ред. А.А. Амангелдиев, Н.Б. Сейсенбек, Н.А. Еламанов. – Қарағанды: «Bilim Innovations Group» орталығы, 2020. – С. 149-155.

Маканова А.У. Планирование целевых индикаторов как часть научной политики Республики Казахстан / The IV International Science Conference «Prospects and achievements in applied and basic sciences» / Под ред. Е. Pluzhnik. – Будапешт, Венгрия, 2021. – С.261-266.

Смагулов К.Е., Кистаубаев Е.Б., Дусекеев Р.М., Ұзақбай С.А., Рамадинқызы К. Авторское свидетельство № 15160 от 15 февраля 2021 г. «Автоматизированная система рейтинговой оценки деятельности сотрудников и структурных подразделений

организации на основе данных научной производительности по базе Scopus».

Смагулов К.Е., Кистаубаев Е.Б., Дусекеев Р.М., Айтмұқаш Д.Б., Ұзақбай С.А., Рамадинқызы К. Авторское свидетельство № 22130 от 3 декабря 2021 г. «Автоматизированная система рейтинговой оценки авторов и организаций на основе данных научной производительности по базе Scopus».

Кудайбергенова Р.Е., Смагулов К.Е., Омиралиева Г.К., Ұзақбай С.А. Наукометрический анализ публикационной активности университетов Казахстана // Вестник КазНУ. Серия педагогические науки. – 2022. – № 3(72).

## **ЗАКЛЮЧЕНИЕ**

В данной монографии представлены результаты исследований в рамках грантового финансирования Комитета науки МОН РК, направленных на дальнейшую цифровизацию процессов в университетах.

Ценность предлагаемого подхода, указанного в первой главе, заключается в сборе на регулярной и непрерывной основе сфокусированной информации о состоянии здоровья студенческой молодежи, имеющей стратегическое значение с точки зрения устойчивого национального развития как основного источника квалифицированного кадрового потенциала. Кроме того, студенты, как и другие категории обучающихся, относятся к группам организованного населения, удобного для успешного проведения целевых оздоровительных мероприятий. Авторы исследования ожидают сформировать научную и технологическую основу для продвижения идеи и принципов «Здоровый Университет – Здоровая нация» среди организаций образования Республики Казахстан, что может быть реализовано путем аккумуляции данных со всех вузов на базе информационного портала КазНУ или путем предоставления технологии другим организациям образования, в том числе на коммерческой основе.

Разработанная интеллектуальная информационно-аналитическая система с визуализацией результатов агрегации и анализа данных о показателях здоровья и показателях, связанных со здоровьем, в том числе, в контексте регионов Казахстана, может быть использована различными группами внутренних и внешних заинтересованных лиц в качестве регулярно обновляющегося информационного ресурса для разработки различных программ поддержки социального или медицинского характера.

Использование международных стандартизованных опросников ВОЗ и международной классификации функционирования, ограничений жизнедеятельности и здоровья детей и подростков

(МКФ, ВОЗ) для кодирования и регистрации данных позволяет обеспечить сопоставимость результатов с соответствующими международными исследованиями и предоставить информационную базу для принятий решений на страновом уровне.

Исследование здоровья студентов в рамках проекта AP09260767 рассматривается как первый этап широкомасштабного исследования, целью которого является долгосрочный мониторинг здоровья студенческой молодежи.

Автоматизированная система рейтинговой оценки авторов и организаций на основе данных научной производительности по базе Scopus предназначена для анализа количественных и качественных показателей ученых и организаций Казахстана, выгружаемых посредством разработанного робота и канала API-Scopus. Система позволяет ранжировать авторов и университеты Казахстана на основании публикационной активности в абсолютных значениях, а также по сумме рейтинговых баллов.

Отличительной характеристикой данной автоматизированной системы является то, что она основана на тех параметрах и индикаторах, которые соответствуют национальному законодательству в области науки, что не могут предложить другие международные базы данных и информационно-аналитические платформы. Так, в системе внедрена возможность отфильтровывать статьи, опубликованные в журналах, индексация которых в базе данных Scopus прекращена; в связи с тем, что в нормативно-правовой базе Казахстана по публикационной активности в рейтинговых изданиях отдельное внимание уделяется основному автору, были введены такие показатели как: публикации автором для корреспонденции, публикации первым автором. Учитывая, что отдельные казахстанские вузы активно участвуют в рейтинге QS, был также введен специализированный фильтр, который, в соответствии с методологией QS, позволяет исключать из учета публикации, в которых авторами указаны 15 и более различных аффилиаций.

Следует отметить, что в современном мире увеличивается значимость производимой учеными научной продукции (публикации, патенты и др.), доступ к которым обеспечивает владение



актуальной научной информацией. В этой связи формирование рейтинга университетов и ученых Казахстана на основе наукометрических показателей позволит выявить лидеров среди научного сообщества, а также организаций, демонстрирующих высокую публикационную активность, оценить вклад ученых в отдельные научные направления, оценку актуальности проводимых исследований на основе публикаций результатов в ведущих рейтинговых изданиях мира и др.

## СПИСОК ЛИТЕРАТУРЫ

1. “Health behavior in school-aged children”, MRC/CSO Social and Public Health Sciences Unit, University of Glasgow [Электронный ресурс]. – URL: <https://hbsc.org/>
2. Нефедовская Л.В., Состояние и проблемы здоровья студенческой молодежи // Литтерра, Москва. – 2017. – стр.192.
3. Баранова А.А., Кучмы В.Р., Физическое развитие детей и подростков Российской Федерации: Сб.мат-лов (выпуск VI) //ПедиатрЪ, Москва – 2013. – стр. 192.
4. “Understanding Generation Z in the workplace”, 2022 Global Automotive Consumer Study [Электронный ресурс]. – URL: <https://www2.deloitte.com/us/en/pages/consumer-business/articles/understanding-generation-z-in-the-workplace.html>
5. Абдрахманова Ш.З., Ахметов В.И., Адаева А.А., Слажнева Т.И., Факторы, формирующие здоровье и детей и подростков Казахстана. Health behavior in school-aged children 2018 // Национальный отчет и Национальный центр общественного здравоохранения, Нур-Султан. – 2019. – стр. 150.
6. Қалиев Т.Б, Қайдарова Ә.С., Қаримова Ж.К., Әшімханова Д.Ә., Маульшариф М.М, Насимова Г.О., Негай Н.А., Сыдықназаров М. Қ., Шаповал Ю.В., Мусатаева Ф.М., Окасова Г.Е., Қасымбеков А.М., Молодежь Казахстана – 2018 и Youth of Kazakhstan – 2018//Национальный доклад и National report, Астана. – 2018. – стр. 410.
7. «Исследование распространенности, глубинных причин и факторов риска и защиты в области суицида и суицидальных попыток в Республике Казахстан» // Детский Фонд ООН (ЮНИСЕФ) в Республике Казахстан, Астана. – 2014. – стр. 108.
8. «Организация экономического сотрудничества и развития (ОЭСР). Краткий обзор образования: показатели ОЭСР» // ОЭСР, Париж – 2017.
9. Lechner K.E., Garcia C.M., Frerich E.A., Lust K., Eisenberg M.E., College students’ sexual health: personal responsibility or the responsibility of the college? [Электронный ресурс]. – 2013. – URL: <http://europepmc.org/abstract/MED/23305542>

10. Hunt J., Eisenberg D., Mental health problems and help-seeking behavior among college students – 2010. – URL: <https://pubmed.ncbi.nlm.nih.gov/20123251/>
11. Quinn P.D., Fromme K., Alcohol use and related problems among college students and their noncollege peers: the competing roles of personality and peer influence [Электронный ресурс]. – 2011. – URL: <https://europepmc.org/article/MED/21683044>
12. «Основы европейской политики и стратегия для XXI века. Здоровье-2020» [Электронный ресурс]. – 2020. – URL: <http://www.euro.who.int/>
13. Patton G.C., Sawyer S.M., Santelli J.S., Ross D.A., Afifi R., Allen N.B., Arora M., Azzopardi P., Baldwin W., Bonell C., Kakuma R., Kennedy E., Mahon J., McGovern T., Ali H. Mokdad, Patel V., Petroni S., Reavley N., Taiwo K., Waldfogel J., Wickremarathne D., Barroso C., Bhutta Z., Fatusi O.A., Mattoo A., Diers J., Fang J., Ferguson J., Ssewamala F., Viner R.M., A Lancet commission on adolescent health and wellbeing [Электронный ресурс]. – 2016. – URL: <https://pubmed.ncbi.nlm.nih.gov/27174304/>
14. Currie C., Социальные детерминанты здоровья и благополучия подростков. Исследование «Поведение детей школьного возраста в отношении здоровья»: международный отчет по результатам обследования 2009–2010 гг. //Европейское региональное бюро ВОЗ Серия «Политика охраны здоровья детей и подростков» – 2012.
15. Barrett M.A., Humblet O., Hiatt R.A., Adler N.E., Big Data and Disease Prevention: From Quantified Self to Quantified Communities [Электронный ресурс]. – 2013. – URL: <https://doi.org/10.1089/big.2013.0027>
16. Закон Республики Казахстан «О государственной молодежной политике» от 9 февраля 2015 года № 285-V ЗПК [Электронный ресурс]. – URL: <https://doi.org/10.1089/big.2013.0027> <http://edu.resurs.kz/elegal/molodezhnaya-politika>
17. О здоровье народа и системе здравоохранения. Кодекс Республики Казахстан от 7 июля 2020 года № 360-VI ЗПК [Электронный ресурс]. – URL: <https://adilet.zan.kz/rus/docs/K2000000360v>
18. Указ Главы государства от 1 июля 2008 г. № 625
19. Catherine N.R., Georgianne F.T., Laura K., Tim M., Shannon L.M., Caitlin L.M., Sarah M.L., Michelle K.B., Frances A., Kathleen A.E.
20. Catherine N.R., Georgianne F.T., Laura K., Tim M., Shannon L.M., Caitlin L.M., Sarah M.L., Michelle K.B., Frances A., Kathleen A.E.,

- Youth Risk Behavior Surveillance System . [Электронный ресурс]. – 2015. – URL: <https://www.cdc.gov/healthyschools/shi/index.htm>
21. [Электронный ресурс]. – 2015. – URL: <https://www.cdc.gov/healthyyouth/data/yrbs/index.htm>
  22. [Электронный ресурс]. – 2015. – URL: <https://nccd.cdc.gov/YouthOnline/>
  23. [Электронный ресурс]. – 2015.– URL: <https://www.cdc.gov/healthyschools/about>
  24. [Электронный ресурс]. – 2015. – URL: <https://www.cdc.gov/healthyschools/about>
  25. Murnaghan D., Morrison W., Griffith E.J., Bell B.L., Duffley L.A., McGarry K., Manske S., Knowledge exchange systems for youth health and chronic disease prevention: a tri-provincial case study // Chronic Diseases and Injuries in Canada [Электронный ресурс]. – 2013. – URL: <https://pubmed.ncbi.nlm.nih.gov/23987222/>
  26. Heinz A., Claire van D., Kern R.M., Catunda C., Helmut erich W., Trends from 2006 - 2018 in Health Behaviour, Health Outcomes and Social Context of Adolescents in Luxembourg [Электронный ресурс]. – 2020. – URL: <https://www.researchgate.net/publication/342630277>.
  27. “Measuring and monitoring children and young people’s mental wellbeing”. Toolkit for schools and colleges. – 2006. – Vol.91(2).
  28. [Электронный ресурс]. – URL: <https://warwick.ac.uk/fac/sci/med/research/platform/wemwbs/about/>
  29. [Электронный ресурс]. – URL: <https://new.boxallprofile.org/#how>
  30. Benjamin S., Constanze K., Mert A., Karim E.A., Literature Review on Medicine Recommender Systems // International Journal of Advanced Computer Science and Applications. – 2019. – P. 82.
  31. Qiwei H., Mengxin J., Inigo Martinez R. T., Manas G., Leid Z., A Hybrid Recommender System for Patient [Электронный ресурс]. – 2018. – URL: <https://www.researchgate.net/publication/330880005>.
  32. Joris F., Amine A.Y., Frédéric M.B.H., Brigitte D., Alpha D., Visual instance-based recommendation system for medical data mining // Procedia Computer Science 112. – 2017. – P. 1747-1754.
  33. Martin W., Daniel P., Concepts, Requirements, Technical Basics and Challenges // International Journal of Environmental Research and Public Health – 2014. – P. 2580 – 2607.
  34. «Анализа положения в области народонаселения в Республике Казахстан» [Электронный ресурс]. – 2019. – URL: <https://kazakhstan.unfpa.org>

35. [Электронный ресурс]. – URL: <https://hbosc.org/>
36. [Электронный ресурс]. – URL: [https://www.who.int/europe/initiatives/health-behaviour-in-school-aged-children-\(hbosc\)-study](https://www.who.int/europe/initiatives/health-behaviour-in-school-aged-children-(hbosc)-study)
37. Jason M. N., New Findings From the Health Behaviour in School-aged Children (HBSC) Survey: Social Media, Social Determinants, and Mental Health. *Adolescent Health*. – 2020. – Vol. 66, No 6.
38. “WONDER online databases utilize a rich ad-hoc query system for the analysis of public health data” [Электронный ресурс]. – URL: <https://wonder.cdc.gov/>
39. “CDC WONDER Online Databases” [Электронный ресурс]. – URL: <https://wonder.cdc.gov/DataSets.html>
40. Friede A., Daniel H. Rosen, Joseph A. R., *Journal of the American Medical Informatics Association* – 1994. – Vol.1(4), – P. 303–312.
41. “Public Health Surveillance, Preparing for the Future” [Электронный ресурс]. – URL: <https://www.cdc.gov/surveillance/pdfs/Surveillance-Series-Bookleth.pdf>
42. “Public Health Surveillance and Data” [Электронный ресурс]. – URL: [https://www.cdc.gov/surveillance/projects/dmi-initiative/where\\_does\\_our\\_data\\_come\\_from.html](https://www.cdc.gov/surveillance/projects/dmi-initiative/where_does_our_data_come_from.html)
43. [Электронный ресурс]. – URL: <https://eupha.org/>
44. “Digital technologies for key public health functions: results of an ECDC expert consultation” [Электронный ресурс]. – URL: <https://www.ecdc.europa.eu/sites/default/files/documents/Digital-technologies-for-public-health-key-functions.pdf>
45. [Электронный ресурс]. – URL: <https://eupha.org/> <https://www.kaznu.kz>
46. [Электронный ресурс]. – URL: [https://www.who.int/health-topics/adolescent-health#tab=tab\\_1](https://www.who.int/health-topics/adolescent-health#tab=tab_1)
47. Robert C., Freeman J., Samda O., Schnohr C. W., de Looze M. E., Nic Gabhainn S., Iannotti R., Rasmussen M., International HBSC Study Group (2009), The Health Behaviour in School-aged Children (HBSC), study: methodological developments and current tensions // *International journal of public health*, – Vol.54(2), – P.140–150. doi:10.1007/s00038-009-5405-9
48. [Электронный ресурс]. – URL: <https://www.who.int/publications/list/2016/icd-children/ru/>
49. [Электронный ресурс]. – URL: <https://apps.who.int/iris/bitstream/handle/10665/85389/924454542X.pdf?sequence=1&isAllowed=y>
50. Hoens T.R., Blanton M., Steele A., Chawla N.V., *Reliable medical*

- recommendation systems with patient privacy // ACM Trans. Intell. Syst. Technol. – 2013. – Vol. 4(4). – P. 1–31.
51. Министерство здравоохранения Российской Федерации. ПРИКАЗ от 1 октября 2020 года N 1053н «Об утверждении стандартов медицинской помощи взрослым при сахарном диабете 1 типа» [Электронный ресурс]. – URL: <https://docs.cntd.ru/document/566406946>
  52. Общество нефрологов, врачей диализа и трансплантологов Казахстана. Международные клинические рекомендации [Электронный ресурс]. – URL: <https://nephro.kz/ru/international-clinical-guidelines.html>. Об утверждении Стандарта организации оказания эндокринологической помощи в Республике Казахстан. Приказ Министра здравоохранения и социального развития Республики Казахстан от 6 июня 2016 года № 478. Зарегистрирован в Министерстве юстиции Республики Казахстан 5 июля 2016 года № 13880 [Электронный ресурс]. – URL: <https://adilet.zan.kz/rus/docs/V1600013880>
  53. [Электронный ресурс]. – URL: [https://apps.who.int/gb/e/e\\_eb142.html](https://apps.who.int/gb/e/e_eb142.html)
  54. [Электронный ресурс]. – URL : [https://apps.who.int/gb/or/e/e\\_wha58r1.html](https://apps.who.int/gb/or/e/e_wha58r1.html)
  55. Hicks, D., Wouters, P., Waltman, L., De Rijcke, S., & Rafols, I. Bibliometrics: The Leiden manifesto for research metrics // Nature. – 2015. - Vol. 520(7548). - P.429-431.
  56. Guozhu M., Xi L., Huibin D., Jian Z., Linyuan W. Way forward for alternative energy research: A bibliometric analysis during 1994–2013 // Renewable and Sustainable Energy Reviews. – 2015. – Vol. 48. – P. 276-286.
  57. Zhou F, Guo HC, Ho YS, Wu CZ. Scientometric analysis of geostatistics using multivariate methods // Scientometrics – 2007. – Vol.73(3). – P.265–79.
  58. Senanayake, U., Piraveenan, M., Zomaya, A. The pagerank-index: Going beyond citation counts in quantifying scientific impact of researchers // PLoS ONE. – 2015. – Vol.10(8). doi:10.1371/journal.pone.0134794
  59. “Data collection”, CWTS Leiden Ranking. Universiteit Leiden Centre for Science and Technology Studies [Электронный ресурс]. URL: <https://www.leidenranking.com/>
  60. Waltman, L. R., Calero Medina, C. M., Kosten, J., Noyons, E. C. M.,

- Tijssen, R. J. W., Eck, N. J. P. van, ... Wouters, P. (2012). The Leiden Ranking 2011/2012: Data collection, indicators, and interpretation. Centre for Science and Technology Studies, Leiden University. Retrieved from <https://hdl.handle.net/1887/19353>.
61. Leydesdorff L., Opthof T., Normalization, CWTS indicators, and the Leiden Rankings: Differences in citation behavior at the level of fields. – 2010.
  62. Marianne G. Counting methods introduced into the bibliometric research literature 1970–2018: A review // *Quantitative Science Studies*. – 2021. – Vol.2(3). – P.932-375. doi.org/10.1162/qss\_a\_00141.
  63. Nikolic, S., Penca, V., Ivanovic, D., Surla, D. Konjovic, Z. Storing of Bibliometric Indicators in CERIF Data Model [Электронный ресурс]. – URL: [https://www.researchgate.net/publication/271015214\\_Storing\\_Of\\_Bibliometric\\_Indicators\\_In\\_Cerif\\_Data\\_Model](https://www.researchgate.net/publication/271015214_Storing_Of_Bibliometric_Indicators_In_Cerif_Data_Model).
  64. Abbott A, Cyranoski D, Jones N, et al. Metrics: Do metrics matter? // *Nature*. - 2010. – Vol. 465, No 2. - P. 860-862.
  65. Kalemis, K., Saba, R. & Elpida, A. The Value of Measuring Intellectual Capital (IC) In Higher Education – A New Challenge of Our Days [Электронный ресурс]. – URL: [https://conferences.ionio.gr/icil2012/download.php?f=papers/186-bou-saba-full\\_text-en-v001.pdf](https://conferences.ionio.gr/icil2012/download.php?f=papers/186-bou-saba-full_text-en-v001.pdf)
  66. Meho, Lokman I., and Cassidy R. Sugimoto. «Assessing the Scholarly Impact of Information Studies: A Tale of Two Citation Databases – Scopus and Web of Science» // *Journal of the American Society for Information Science and Technology*. – 2009. – Vol.60(12). – P.2499–2508.
  67. Haddow, G., Genoni, P. Citation analysis and peer ranking of Australian social science journals // *Scientometrics*. – 2010. – Vol.85(2), – P.471-487.; Kulkarni, A. V., Aziz, B., Shams, I., & Busse, J. W. Comparisons of citations in Web of Science, Scopus, and Google Scholar for articles published in general medical journals // *JAMA*. – 2009. – Vol.302(10), – P.1092-1096.; Torres-Salinas, D., Lopez-Cózar, E. D., Jimenez-Contreras, E. Ranking of departments and researchers within a university using two different databases: Web of Science versus Scopus // *Scientometrics*. – 2009. – Vol.80(3), – P.761-774.
  68. Mongeon, P., & Paul-Hus, A. (2016). The journal coverage of Web of Science and Scopus: A comparative analysis // *Scientometrics*. – 2016. – Vol.106(1), – P. 213-228.

69. Abramo G. Revisiting the scientometric conceptualization of impact and its measurement // *Journal of Informetrics*. – 2018. – Vol.12(3), – P.590–597.
70. Gupta, H.M., Campanha, J.R., Pesce, R. Power-law distributions for the citation index of scientific publications and scientists // *Brazilian Journal of Physics*. – 2005. – Vol.35(4 A), – P.981-986. doi:10.1590/S0103-9733200500060001.
71. Dag W. Aksnes, Liv Langfeldt, Paul Wouters. Citations, Citation Indicators, and Research Quality: An Overview of Basic Concepts and Theories // *SAGE Open*. – 2019. – Vol.1(7). DOI: 10.1177/2158244019829575.
72. Aksnes, D. W. A macro study of self-citation // *Scientometrics*. – 2003. – Vol. 56(2). – P. 235-246.
73. Fowler, J. H., Aksnes, D. W. (2007). Does self-citation pay // *Scientometrics*. – 2007. – Vol.72(3), – P. 427–437. <https://doi.org/10.1007/s11192-007-1777-2>.
74. Hirsch, J. E. An index to quantify an individual's scientific research output // *Proceedings of the National Academy of Sciences of the United States of America*. – 2005. - Vol. 702, No 46. – P. 16569-16572.
75. Lippi G, Borghi L. A short story on how the H-index may change the fate of scientists and scientific publishing // *Clinical Chemistry and Laboratory Medicine (CCLM)*. – 2014. – Vol.52. – P.1-3. <https://doi.org/10.1515/cclm-2013-0715>.
76. Lippi, G., Mattiuzzi, C. Scientist impact factor (SIF): A new metric for improving scientists' evaluation? // *Annals of Translational Medicine*. – 2017. – Vol.5(15). doi:10.21037/atm.2017.06.24.
77. Gauffriau, M., Larsen, P. O. Counting methods are decisive for rankings based on publication and citation studies // *Scientometrics*. – 2005. – Vol. 64, No 1. - P. 85-93.
78. Vavrycuk V. Fair ranking of researchers and research teams [Электронный ресурс]. - 2018. – URL: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0195509>.
79. Hagen, N. T. Harmonic allocation of authorship credit: Source-level correction of bibliometric bias assures accurate publication and citation analysis [Электронный ресурс]. - 2008. – URL: doi:10.1371/journal.pone.0004021
80. Vavrycuk V. Fair ranking of researchers and research teams [Электронный ресурс]. - 2018. – URL: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0195509>



81. Gauffriau, M., Larsen, P. O., Maye, I., Roulin-Perriard, A., Von Ins, M. Publications, cooperation and productivity measures in scientific research // *Scientometrics*. – 2007. – Vol. 73, No 2. - P. 175-214.
82. Lindsey, D. Production and citation measures in the sociology of science: The problem of multiple authorship // *Social Studies of Science*. – 1980. –Vol. 10, No 2. - P. 145-162. Lange, L. L. Citation counts of multi-authored papers—First-named authors and further authors // *Scientometrics*. – 2001. – Vol. 52, No 3. - P. 457-470.
83. Moya-Anegón, F., Guerrero-Bote, V.P., Bornmann, L., Moed, H.F. The research guarantors of scientific papers and the output counting: A promising new approach // *Scientometrics*. – 2013. –Vol. 97, No 2. - P. 421-434. Huang, M. H., Lin, C. S., & Chen, D. Z. Counting methods, country rank changes, and counting inflation in the assessment of national research productivity and impact // *Journal of the American Society for Information Science and Technology*. – 2011. –Vol. 62, No 12. - P. 2427-2436. Lin, C. S., Huang, M. H., Chen, D. Z. The influences of counting methods on university rankings based on paper count and citation count // *Journal of Informetrics*. - 2013. – Vol. 7, No 3. - P.611-621.
84. Abramo, G., D'Angelo, C. A., Rosati, F. The importance of accounting for the number of co-authors and their order when assessing research performance at the individual level in the life sciences // *Journal of Informetrics*. – 2013. – Vol. 7, No 1. - P. 198-208.
85. SCImago Research Group. Description of SCImago Journal Rank Indicator. Accessed 2012. URL: <http://www.scimagojr.com/SCImagoJournalRank.pdf> and <http://arxiv.org/ftp/arxiv/papers/0912/0912.4141.pdf>
86. Waltman, L., & Van Eck, N. J. Field-normalized citation impact indicators and the choice of an appropriate counting method // *Journal of Informetrics*. – 2015. – Vol. 9, No 4. - P. 872-894.
87. Marusić, A., Bosnjak, L., Jerončić, A. A systematic review of research on the meaning, ethics and practices of authorship across scholarly disciplines [Электронный ресурс]. - 2011. – URL: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0023477>
88. Frandsen, T. F., & Nicolaisen, J. What is in a name? Credit assignment practices in different disciplines // *Journal of Informetrics*. – 2010. –Vol. 4, No 4. - P. 608-617.
89. Waltman, L., & Van Eck, N. J. Field-normalized citation impact indicators and the choice of an appropriate counting method //

- Journal of Informetrics. – 2015. – Vol. 9, No 4. - P. 872-894.
90. Lin, C., Huang, M., Chen, D. The influences of counting methods on university rankings based on paper count and citation count. *Journal of Informetrics*. – 2013. – Vol.7(3), – P.611-621. doi:10.1016/j.joi.2013.03.007
  91. Gauffriau, M., Larsen, P. O., Maye, I., Roulin-Perriard, A., Von Ins, M. Comparisons of results of publication counting using different methods. *Scientometrics*. – 2008. – Vol.7(1), – P.147-176.
  92. Hagen, N. T. Harmonic allocation of authorship credit: Source-level correction of bibliometric bias assures accurate publication and citation analysis [Электронный ресурс]. - 2008. – URL: doi:10.1371/journal.pone.0004021
  93. Abbas, A. M. Weighted indices for evaluating the quality of research with multiple authorship // *Scientometrics*. – 2011. - Vol. 55, No 1. - P. 107-131
  94. Egghe, L., Rousseau, R., & Van Hooydonk, G. Methods for accrediting publications to authors or countries: Consequences for evaluation studies // *Journal of the American Society for Information Science*. – 2000. - Vol. 51, No 2. - P. 145-157.
  95. Assimakis, N., Adam, M. (2010). A new author's productivity index: p-index // *Scientometrics*. - 2010. - Vol. 85, No 2. - P. 415-427.
  96. Stallings, J., Vance, E., Yang, J., Vannier, M. W., Liang, J., Pang, L., & Wang, G. Determining scientific impact using a collaboration index / *Proceedings of the National Academy of Sciences of the United States of America*. – 2013. - Vol. 110, No 24. - P. 9680-9685.
  97. Hagen, N. T. Harmonic publication and citation counting: sharing authorship credit equitably-not equally, geometrically or arithmetically // *Scientometrics*. – 2010. –Vol. 84, No 3. - P. 785-793.
  98. Hagen, N.T. Harmonic publication and citation counting: Sharing authorship credit equitably - not equally, geometrically or arithmetically. *Scientometric*. – 2010. – Vol.84(3), – P.785-793. doi:10.1007/s11192-009-0129-4.
  99. Abbas, A. M. Weighted indices for evaluating the quality of research with multiple authorship. *Scientometrics*. – 2011. – Vol.55(1), – P.107-131.
  100. Egghe, L., Rousseau, R., Van Hooydonk, G. Methods for accrediting publications to authors or countries: Consequences for evaluation studies. *Journal of the American Society for Information Science*. – 2000. – Vol.51(2), – P.145-157.

101. Assimakis, N., Adam, M. A new author's productivity index: P-index. *Scientometrics*. – 2010. – Vol.85(2), - P. 415-427. doi:10.1007/s11192-010-0255-z
102. Kosmulski, M. The order in the lists of authors in multi-author papers revisited // *Journal of Informetrics*. – 2012. - Vol. 6, No 4. - P. 639-644.
103. Aziz, N. A., & Rozing, M. P. Profit (p)-index: The degree to which authors profit from co-authors [Электронный ресурс]. - 2013. – URL: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0059814>
104. Kotsemir, M., & Shashnov, S. (2017). Measuring, analysis and visualization of research capacity of university at the level of departments and staff members. *Scientometrics*, 112(3), 1659-1689. doi:10.1007/s11192-017-2450-7
105. Lippi G, Borghi L. A short story on how the H-index may change the fate of scientists and scientific publishing. *Clin Chem Lab Med* 2014;52:e1-3.
106. Lippi G, Mattiuzzi C. The challenges of evaluating scientists by H-index and citations in different biomedical research platforms. *Clin Chim Acta* 2013;421:57
107. Lippi, G., & Mattiuzzi, C. (2017). Scientist impact factor (SIF): A new metric for improving scientists' evaluation? *Annals of Translational Medicine*, 5(15) doi:10.21037/atm.2017.06.24

## СОДЕРЖАНИЕ

Введение .....	3
1. ИНТЕЛЛЕКТУАЛЬНАЯ ИНФОРМАЦИОННО-АНАЛИТИЧЕСКАЯ СИСТЕМА «ЦИФРОВОЙ ПРОФИЛЬ ЗДОРОВЬЯ СТУДЕНТА».....	5
1.1 Анализ существующих решений по оценке состояния здоровья молодого поколения.....	6
1.1.1 Проблемы оценки состояния здоровья молодежи в мире и РК.....	7
1.1.2 Цифровые технологии для оценки состояния здоровья молодежи .....	10
1.1.3 Применение опросников для сбора медицинских данных .....	12
1.1.4 Цифровые платформы по контролю и профилактике заболеваний молодежи .....	17
1.2 Техническое описание интеллектуальной информационно-аналитической системы.....	26
1.2.1 Источники данных .....	26
1.2.2 Организация хранилищ данных .....	30
1.2.3 Модуль анкетирования студентов .....	33
1.2.4 Модуль интеллектуального анализа данных .....	35
1.2.5 Модуль визуализации данных .....	37
1.2.6 Модуль поддержки принятия решений и выработки рекомендаций .....	47
1.2.7 Обеспечение безопасности и целостности данных.....	52
1.3 Структура интеллектуальной информационно-аналитической системы и прототип веб-приложения.....	54
1.4 Список статей, опубликованных в рамках исследования .....	63

2. АВТОМАТИЗИРОВАННАЯ СИСТЕМА ПО УЧЕТУ ПУБЛИКАЦИОННОЙ АКТИВНОСТИ УЧЕНЫХ И ОРГАНИЗАЦИЙ КАЗАХСТАНА МЕТОДОМ БИБЛИОМЕТРИЧЕСКОГО АНАЛИЗА НА ОСНОВЕ ДАННЫХ API SCOPUS .....	65
2.1 Анализ библиометрических методов оценки эффективности публикационной активности авторов и организаций .....	66
2.2 Техническое описание автоматизированной системы по оценке публикационной активности на основе данных из базы Scopus .....	78
2.3 Структура автоматизированной системы по оценке публикационной активности ученых и организаций Казахстана на основе рейтингового и балльного подхода ...	94
2.3.1 Рейтинговая оценка авторов.....	95
2.3.2 Рейтинг Авторов – балловая система.....	105
2.3.3 Рейтинговая оценка организаций .....	106
2.3.4 Рейтинг Организаций – балловая система.....	114
2.4 Список статей, опубликованных в рамках исследования.....	115
Заключение .....	118
Список литературы .....	121

Научное издание

**ЦИФРОВЫЕ РЕШЕНИЯ ДЛЯ ПОВЫШЕНИЯ  
КАЧЕСТВА УПРАВЛЕНИЯ УНИВЕРСИТЕТОМ**

*Монография*

*Под общей редакцией К.Е. Смагулова и М.Е. Мансуровой*

Выпускающий редактор *Г. Бекбердиева*  
Компьютерная верстка и дизайн обложки *А.Калиева*

**ИБ № 8**

Подписано в печать 26.10.2022. Формат 64x80 <sup>1</sup>/<sub>16</sub>. Бумага офсетная.

Печать цифровая. Объем 8,3. Тираж 500 экз. Заказ №8.

Издательский дом «Smart University Press».

050040, г. Алматы, ул. Кенесары хана, 54/43.

Отпечатано в издательского дома «Smart University Press»



## How to Use This Book

This book is designed to help you learn the fundamentals of computer science and programming. It is divided into several parts, each covering a different aspect of the subject. The first part, "Introduction to Computer Science," provides a general overview of the field and its history. The second part, "Programming Fundamentals," covers the basic concepts and syntax of programming. The third part, "Data Structures and Algorithms," discusses the ways in which data is organized and processed. The fourth part, "Computer Systems and Architecture," explores the hardware and software components of a computer system. The fifth part, "Applications and Case Studies," provides examples of how computer science is used in various fields. The final part, "Conclusion and Future Trends," discusses the current state of the field and its potential future developments.

The book is written in a clear and concise style, with a focus on practical examples and exercises. It is suitable for students and professionals alike who are interested in learning more about computer science and programming. The book is available in both print and digital formats, and can be accessed through various online platforms.

For more information about this book, please visit our website at [www.example.com](http://www.example.com). We would be happy to answer any questions you may have.

Thank you for your interest in this book. We hope you find it helpful and enjoyable to read.

Best regards,  
The Authors

© 2023 by the authors. All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, without the prior written permission of the publisher.

This book is published by Example Publisher, 123 Main Street, New York, NY 10001. For more information, please contact our sales department at [sales@example.com](mailto:sales@example.com).

The cover image is a stock photo by Example Photographer. All other images are the property of their respective owners.

The authors would like to thank their families and friends for their support and encouragement throughout the writing process.

The authors would also like to thank the reviewers for their helpful comments and suggestions.

The authors would like to thank the publisher for their support and for making this book possible.

The authors would like to thank the readers for their interest in this book and for their feedback.

The authors would like to thank the world for the opportunity to share their knowledge and experiences.

The authors would like to thank the future for the potential of computer science and programming.

The authors would like to thank the past for the lessons learned and the experiences gained.

The authors would like to thank the present for the challenges and opportunities that lie ahead.