***ФГБОУ ВО «Кубанский государственный аграрный университет имени И.Т. Трубилина»,***

***Российская Федерация***

***Ватамов Д.Н. ПИ2104***

**153 Теория информации, данные и знания. Практическая 2. Инструкция по разработке собственного приложения 2020-10-22**

## Заголовок

Разработка и публикация приложения: Инструкция, требования, регистрация в РИНЦ/ResearchGate.

## Резюме текста

**1. Введение и постановка задачи**
Преподаватель проводит практическое занятие №2 для группы ИТ-2003 по дисциплине "Теория информации, данные, знания". Основная задача занятия – изучение инструкции по разработке собственного приложения и выбор темы для него.

**2. Важность задания и контекст дистанционного обучения**
Подчеркивается важность выполнения задания для успешной сдачи экзамена ("сам экзамен"). Преподаватель выражает обеспокоенность снижением качества знаний при дистанционном обучении, ссылаясь на исследования (в т.ч. в США) и возможное негативное отношение работодателей к выпускникам, учившимся удаленно (потенциальное снижение зарплаты на 12-13%). Отмечается недостаток обратной связи и контроля при дистанционном формате по сравнению с очным, где преподаватель видит работу каждого студента. Обсуждается гипотетическая возможность улучшения ситуации с помощью VR-технологий в будущем.

**3. Суть практического задания: Разработка и публикация приложения**
Студенты должны разработать собственное приложение, используя систему AIDOS и полученные знания (обработка табличных данных, текстов, графики, риэлторская оценка). Цель – переход от знаний к умениям и навыкам. Задание включает:

* Выбор уникальной темы (не повторяющейся с уже существующими 208 приложениями в AIDOS-облаке).
* Создание приложения.
* Написание описания работы по предложенным шаблонам (статьям преподавателя).
* Размещение приложения и его описания в научных базах данных: ResearchGate и РИНЦ (Российский индекс научного цитирования).

**4. Требования к теме и данным**

* **Уникальность:** Тема не должна дублировать существующие в AIDOS-облаке.
* **Соответствие законодательству и этике:** Тема не должна противоречить законам РФ и морально-этическим нормам.
* **Научная значимость:** Тема должна представлять собой научное исследование, решающее задачу с научной или практической ценностью.
* **Источники данных:** Можно использовать данные из репозиториев (UCI, Kaggle), официальной статистики (Росстат, Генпрокуратура), найти самостоятельно или использовать данные из предыдущих работ (но с более качественным описанием). Рекомендуется выбирать задачи с числовыми данными и простой структурой (1-2 файла CSV).

**5. Регистрация в РИНЦ и ResearchGate**

* **РИНЦ:** Требуется регистрация на сайте elibrary.ru и обязательная регистрация в Science Index (поставить галочку при заполнении анкеты) для получения SPIN-кода. Процесс может занять несколько недель. После получения SPIN-кода заключается договор на физлицо для размещения непериодических изданий (статей, монографий).
* **ResearchGate:** Регистрация проще, но требует корпоративного адреса электронной почты (домен @kubsa.ru или @edu.kubsa.ru). Студентам необходимо выяснить в университете (ЦИТ, деканат, куратор), как получить такой адрес или подтверждение для регистрации.

**6. Оценка и соавторство**

* **Оценка:** Зависит от уровня публикации:
	+ Отлично: Разработка + размещение в AIDOS-облаке, ResearchGate и РИНЦ.
	+ Хорошо: Разработка + размещение в AIDOS-облаке и ResearchGate.
	+ Удовлетворительно: Разработка + размещение в AIDOS-облаке (при условии достаточного качества работы).
	+ Неудовлетворительно ("двойка"): Отсутствие выполненного задания.
* **Соавторство:** Если студент не может самостоятельно разместить работу в РИНЦ/ResearchGate, преподаватель может помочь с размещением, но будет указан как соавтор (на что имеет основания как разработчик технологии и методики). Преподаватель предпочел бы, чтобы студенты делали это сами, так как у него уже большое портфолио (более 630 работ).

**7. Возможность коллаборации**
Задание можно выполнять вдвоем, но не более. В этом случае оба студента могут получить "сам экзамен".

**8. Демонстрация работы с данными (CSV)**
Преподаватель показывает, как импортировать CSV-файлы в систему AIDOS, настраивать параметры (разделитель, кодировка) и запускать анализ для построения модели и оценки ее достоверности. Отмечается, что система AIDOS способна обрабатывать большие объемы данных (миллионы записей).

**9. Заключение и дальнейшие шаги**
Студентам необходимо немедленно начать процесс регистрации в РИНЦ и ResearchGate, выбрать тему и приступить к поиску данных и разработке приложения. Преподаватель готов консультировать по возникающим вопросам.

## Детальная расшифровка текста

**1. Введение и постановка задачи**

* **Приветствие и организационные моменты**
Здравствуйте, ребята.
Здравствуйте.
Здравствуйте, здравствуйте.
Ну что? Сейчас мы без десяти отметимся в двух системах. В одной я себя отмечаю, в другой вас. И начнём занятие.
* **Определение темы и плана занятия**
Сегодня у нас 22 октября 2020 года. Четвёртая пара, 13:50-15:20. Практическое занятие номер два с группой ИТ-2003 по дисциплине Теория информации, данные, знания. И мы на этой работе, на этом практическом занятии изучаем инструкцию по разработке собственного приложения. Ищем себе тему.
Значит, давайте сейчас я сделаю совместный доступ к экрану и начнём изучать этот вопрос.
* **Навигация по сайту преподавателя**
На мой сайт выходим. Второй пункт кликаем на моём сайте. И переходим, вот здесь где картинка, здесь внизу написано: задание, инструкция для учащихся по разработке собственного приложения.
* **Проверка предыдущего материала**
И смотрим, была ли у нас эта тема у этой группы? Нет, ещё не было. Тогда, значит, мы начинаем с нуля изучение этого вопроса.
Сейчас в чат, ребята, шлю вам ссылочку на эту инструкцию. И открываю у себя её на экране.

**2. Важность задания и контекст дистанционного обучения**

* **Акцент на важности темы**
Значит, прежде всего, вот сейчас надо вам внимательно-внимательно меня послушать, прямо вот обратить внимание на то, что я буду говорить.
И сейчас я посмотрю, кто у нас тут староста. Вот у нас 2003 - это Волков, да? Евгений, да?
Да-да-да.
Как-то очень тихо слышно вас. Исправился? Ну ладно, ну, в общем, понять-то можно. Значит, Евгений, это надо передать всем остальным, кого сейчас нет на занятии. Потому что у вас тут побольше там в группе, чем присутствует. Поэтому вот тема сейчас важная для того, чтобы хорошо сдать экзамен по дисциплине. Это о том, как сдать экзамен, понимаете?
* **Проблемы дистанционного обучения и его последствия**
Поэтому, значит, сообщаю вам такую информацию, не очень приятную. Несколько дней назад в одной из новостей, по-моему, в новостях, по-моему, передали, что измерили уровень знаний тех, кто учился дистанционно, с теми, кто учился в прошлом году по этим же дисциплинам, ну, в вузах, в США это было, в США. Вот, в вузах, кто в прошлом году учился, и сейчас. И тех, кто сейчас учился дистанционно, их знания сравнили с теми остаточными знаниями тех, кто учился этому же самому год назад очно. И обнаружили очень нехорошие результаты, что уровень подготовки намного снизился.
И сейчас работодатели говорят о том, что людей, кто учился дистанционно, они будут стараться не брать на работу. И если даже и будут брать, то будут говорить: "Ну вы ж дистанционно учились, поэтому зарплата пониже будет у вас". И они посчитали, что в среднем 12-13% будет зарплата ниже у тех, кто учился дистанционно, чем у тех, кто нормально учился, в нормальных условиях, когда ещё не было этой пандемии.
Вот такие вот грустные новости.
* **Сравнение очного и дистанционного форматов**
Вот, то есть, всё-таки, когда преподаватель стоит там в аудитории, объясняет вам очно всё, индивидуально, подходит к каждому. Я, например, когда в аудитории стою, я вижу все компьютеры, всех студентов. Сейчас я ни одного студента не вижу и не вижу, работаете вы на компьютерах или нет. И если вы, допустим, даже и работаете, я не вижу, что вы там делаете. То есть современные системы дистанционного обучения, они очень ущербно дают возможность преподавания по сравнению с очным.
* **Перспективы VR и проблема обратной связи**
Может быть, когда-нибудь, я не исключаю, может быть, там лет через 10, я одену очки виртуальной реальности и буду находиться в виртуальной аудитории, вы будете тоже находиться в этой же виртуальной аудитории и будете работать за виртуальными компьютерами, а я буду всё это видеть, что вы там делаете. И буду видеть ваши аватары, вы будете мой аватар видеть. Ну, в общем, короче, что-нибудь такое, может быть, и будет, но пока что до этого далеко. И преподаватели все в один голос говорят, что не хватает информации обратной связи.
* **Проверка присутствия и внимания студентов**
То есть я сейчас вот, допустим, вам рассказываю, я не знаю, там вы, не там вы. Чтобы мне это проверить, мне нужно дать вам задание какое-нибудь. Ну, например, возьмите сейчас в чате, поставьте плюсики те, кто слышал, что я сейчас говорил. Ну сейчас поставьте.
(Пауза)
У меня чат не отправляется, но я слышал.
(Пауза)
Ну хорошо. Вот. Ну а теперь, кто понял, тоже плюсики, и уже они идут на сам экзамен. Но это я подшучиваю, ребят, понимаете? То есть вот сейчас вы слышали, ну слава богу. А я вот такие проверки проводил: кто меня слышит, кто понял, что я говорю, значит, вернее, просто, кто меня слышал сейчас, поставить плюсики. И вот семь человек у меня присутствует, а из них там пять поставили плюсики, например, или четыре. А остальные? А остальных ещё семь вообще нету в списке. А те, кто есть, тоже никак не реагируют на то, что я говорю, понимаете? То есть вообще интересно просто. То есть я так подозреваю, что там, может быть, их и нету вообще.

**3. Суть практического задания: Разработка и публикация приложения**

* **Обзор инструкции и цели задания**
Ну ладно. Значит, сейчас вы должны видеть эту инструкцию. Я вам ссылочку послал на неё. Значит, обратите внимание, вот здесь вот есть пункт девять. Здесь написано, видите, что вы должны сделать. Имейте, значит, в виду, что на лабораторных, практических занятиях вы должны развивать навыки. То есть на лекционных занятиях я даю вам знания, а на практических, лабораторных занятиях вы должны эти знания учиться применять, то есть развивать умение применять знания. А потом эти умения должны доводить до уровня навыков, когда уже до определённой степени автоматизма они доводятся, понимаете?
Вот. То есть уже особо так не задумываясь, вы можете применять эти умения на практике.
* **Отличие от стандартных лабораторных работ**
Значит, это называется навыки. Так вот, чтобы выработать эти навыки, вам нужно решить несколько задач. На большинстве лабораторных работ и практических заданий при очном обучении вы выполняли бы сами самостоятельно при мне, под моим наблюдением, на в аудитории, выполняли бы эти задания. А сейчас я вам лабораторные работы показываю на своём компьютере, а вы смотрите на монитор. А теперь приходит момент такой, когда вам надо будет самим что-то делать. Ну, при этом, если у вас будут возникать вопросы, то я вам помогу. То есть я на все ваши вопросы отвечу. Ну, практически на все, которые по дисциплине, по крайней мере.
* **Обзор пройденных тем и возможностей системы AIDOS**
Значит, смотрите, что вам нужно сделать. Мы с вами прошли лабораторные работы разных типов. И простенькую работу проходили, на которой я очень подробно всё объяснял: обработка табличных данных, да, экселевских, в интеллектуальной системе. И интеллектуальная обработка текстов. А потом более сложную уже, такую довольно реалистичную работу проходили по обработке табличных данных, разработка методики риэлторской оценки. А потом я вам даже показал, как обрабатывать графическую информацию, один из вариантов. В системе AIDOS есть много вариантов обработки графической информации. Я вам один из них показал.
* **Этапы выполнения задания**
Вот. И, значит, таким образом, у вас должно возникнуть уже представление о том, что можно делать. И вы должны теперь решить следующий вопрос: вы должны выбрать тему собственного приложения и потом его создать, это приложение. Если будут вопросы, то вы меня спрашиваете, я буду подсказывать, как это делается методически. Вот, если будут возникать вопросы. Создать это приложение, описать его по шаблону. Тут есть четыре шаблона описания, ну это мои статьи и работы. А потом разместить это приложение в ResearchGate и в РИНЦ.

**4. Требования к теме и данным**

* **Уникальность темы**
Теперь, как мы выбираем тему? Ещё одно требование к теме. Значит, эта тема не должна уже, не должна уже быть размещённой в AIDOS-облаке, то есть не должно быть работы на эту тему. Вот здесь у нас в пункте четыре, смотрите, в пункте четыре, у нас указано, видите, что уже имеются AIDOS-приложения. И вот я кликаю вот по этой ссылке, и получается у меня списочек, в котором на данный момент у нас 208 AIDOS-приложений. Почему-то у меня здесь высвечивается 207, но у вас должно 208 высветиться, появиться. Потому что у меня просто кэш тут застрял, а 207 и всё, просто треснет. Надо всё это стирать, чтобы 208. Вот 208 приложений тут.
* **История создания приложений и требование уникальности**
Значит, эти приложения созданы за несколько лет. Вот здесь справа есть дата, когда оно там размещено. И мы видим, что это вот у нас технология размещения, вообще AIDOS-облако вот это, оно появилось в 2017 году, в начале 2017 года. И вот уже 3 года там размещаются приложения. Эти приложения разрабатываются или просто мной, или мной с соавторами, или студентами, или с соавторами, и потом я их размещаю.
* **Требования к соответствию законодательству и этике**
А какие темы нужно затрагивать, чтобы разместить работу в РИНЦ? Не не РИНЗ, а РИНЦ. Российский индекс научного цитирования. Темы могут быть любые, связанные с любым направлением науки. Абсолютно. Тут ограничения нет. Но сейчас я как раз этот вопрос буду обсуждать, как выбрать тему, поэтому чуть-чуть подождите. Ну, в общем, я могу вам так сказать, Марк, что тема не должна противоречить действующему законодательству Российской Федерации и общепризнанным в России морально-этическим нормам.
* **Требование научной значимости**
В остальном может быть любая тема абсолютно. Ну, конечно, эта тема должна быть научного исследования. Что такое тема научного исследования, я вам объясню, если, значит, до этого дойдёт сейчас, дойдём до этого, я вам объясню. То есть вы должны решать какую-то задачу, имеющую научное и практическое значение.
* **Источники данных: UCI, Kaggle, Росстат и другие**
Теперь, откуда мы берём данные? Раньше когда-то, ну, скажем, лет 10 назад, если бы я сделал запрос в интернете, данные для машинного обучения, то было бы там найдено очень мало сайтов, буквально по пальцам пересчитать. Ну, в принципе, вот один сайт такой и есть, это сайт UCI. Вот. Сейчас я покажу. Вот это сайт UCI. Я на них дал ссылочки. Здесь довольно много наборов данных. Набор данных называется dataset. Вот. То есть, ну, достаточно. То есть выбираете какой-то набор данных, который вам нравится, и, значит, смотрите, скачать данные, папка данных, набор данных, описание и сам сами данные.
Вот появился сайт Kaggle. Он появился гораздо позже, чем UCI, он появился лет пять назад, может, если так, если не ошибаюсь, примерно. Вот. Сначала это была самостоятельная компания. Потом этот сайт стал быстро очень лидером мировым в своей области, в области э-э, размещения, скажем так, конкурсов по решению задач, э-э, которые решаются с применением систем искусственного интеллекта. Вот так примерно. Это профиль этого сайта Kaggle. И он очень быстро стал очень популярным. Почему? А потому что он не просто данные размещает, а он организует конкурсы. И эти конкурсы имеют призовые фонды. И бывают фонды такие, я бы сказал, знаете, вот представьте себе, фонд призовой 100.000 долларов, ребят. 100.000 долларов, понимаете? Занимаешь первое место, получаешь половину этой суммы. А те, кто занял второе и третье место, занимают, вторую половину получают вторую половину этой суммы, получают в таком, такой пропорции: тот, кто занял второе место, получает половину от этой оставшейся части, от 50.000. А те, кто занял третье место, они получают, э-э, то есть, вернее, 2/3 получает тот, кто занял второе место от оставшейся части, 1/3 получают те, кто занял третье место от от половины. Значит, есть, ребята, э-э, гроссмейстеры там, мастера, рейтинги оцениваются участников. Есть люди, которые побеждали вот в этих конкурсах, и не однажды. Есть, которые вообще так систематически побеждают в этих конкурсах. И получают вот эти призовые фонды, кстати. Вот. Я интересовался этим, смотрел, как там, как это выглядит, и кто там занимает эти призовые места. Текущие конкурсы. То есть есть смысл разобраться. Я скажу, что вообще это, э-э, ну, скажем так, увлекает, увлекает. То есть если бы вы владели бы технологией, ну, наверное, более совершенной, чем система AIDOS, владели бы какой-то технологией интеллектуальной, языком программирования, то не исключено, что вы могли бы там апробировать свои силы в этом Кагле, и, может быть, там подняться повыше, а может быть, и занять призовое место. И, в общем-то, даже и заработать. Почему бы и нет?
Вот. То есть мы сейчас, я вам рассказал про источники данных основные для решения задач в области искусственного интеллекта. Теперь дополнительные, смотрите. Мы можем взять любой набор данных, который мы найдём просто поиском, поиском просто, данные для машинного обучения. Сейчас таких файлов, таких сайтов с такими данными накопилось уже очень много. Раньше когда-то их вообще не было. Был один UCI только. Вот. Теперь смотрим на мою страничку вот эту вот, lci.kubagro.ru/aidos/P14. А здесь, видите, всякие есть мои разработки, учебные пособия там, статьи. А потом идут шаблоны, но это устаревшая страничка немножко. Здесь шаблоны устаревшие, 2016 год. Нет, самые новые - 2019. А потом указаны, ребята, вот эти вот ссылки на эти сайты с источниками данных. А потом написано, смотрите, 50 лучших баз данных для машинного обучения в публичном доступе. А также, смотрите, ребята, здесь есть ссылочки на работы по, на источники данных по информационной безопасности, сайт Федеральной статистики, ФЭД, сайт прокуратуры Федеральной. Что это такое вообще? Значит, я вам могу сказать, что вы можете поставить вопрос о том, как связано, допустим, развитие тяжёлой промышленности с экологическим состоянием среды, окружающей, природной. И вы можете вот здесь вот в Федеральной статистике по всем регионам России, в разрезе по всем регионам России, посмотреть это. И скачать данные отсюда. И использовать их в работе. Я такие работы проводил. Сейчас я вам покажу, не отходя от кассы. Вот. Экология. Работы по экологии. Работы по экологии основаны на том, что я просто взял и извлёк из статистических баз данных Госстатистики материалы по экологии и по другим аспектам. Ну, например, по качеству жизни. И сравнил. Система AIDOS что делает? Вы, наверное, уже поняли, что она делает. Она выявляет зависимости между какими-то факторами и результатом их влияния на объект. Вот. И вот среди факторов могут быть факторы антропогенные, то есть те, которые связаны с деятельностью людей. А их влияние - это экология среды и качество жизни. Ну, скажем, заболеваемость определёнными группами заболеваний. Или просто заболеваемость всеми статистика по всем заболеваниям. Я вот такие работы проводил. Получается очень неплохо. Оказывается, вполне такие явные есть зависимости, которые можно выявить и исследовать эту моделируемую предметную область, увидеть там какую-то картину, увидеть.
* **Рекомендации по выбору данных**
Я вам советую, ребята, если вы посмотрим на эти вот датасетс, на сами вот эти источники данных, то я могу вам сказать, что есть задачи сложные, где много CSV файлов. А есть задачи простенькие. Ну вот, скажем, два CSV файла. И размер маленький у них, видите, 42, 43 МБ. То есть можно попытаться понять, что они там хотят, прочитать. Вот. И если исходные данные, значит, у них обычно самый простой вариант - это два файла. Один вариант - тренировочная выборка или обучающая выборка, а другой файл - это тестовая выборка для проверки качества решения, для проверки качества работы вашей модели. Нам для наших целей, вот для целей обучения, для того, чтобы сделать задачу, описать её по шаблону, разместить её в AIDOS-облаке, разместить её в ResearchGate и РИНЦе, нам это, в общем-то, нет необходимости в этой тестовой выборке. Нам достаточно тренировочной выборки. Но я могу вам сказать, что, может быть, кто-нибудь из вас заинтересуется этим всем.

**5. Регистрация в РИНЦ и ResearchGate**

* **Процедура регистрации в РИНЦ и Science Index**
Для того, чтобы разместить в РИНЦ и, для этого надо в РИНЦ зарегистрироваться вам и в системе Science Index. Это восьмой пункт, здесь об этом написано. В РИНЦ гораздо сложнее. Ну там как сложнее? Там заполнение формы стандартное, как при регистрации в любом сервисе, почте там, где угодно. Вот, в почтовом сервисе, реформе, абсолютно стандартная регистрация, но там проблема в чём? Они довольно долго там соображают, зарегистрировать вас или нет. То есть это длительное время, понимаете, проходит. В конце концов, значит, вас там регистрируют в системе РИНЦ и Science Index. И после этого вы получаете спин-код. И после этого вы получаете право заключения договора на физическое лицо на размещение непериодических изданий. И можете сами в РИНЦ размещать свои публикации, которые размещены в ResearchGate. Такие публикации, они считаются публикациями с присвоением DOI. И, в общем, это спокойно их размещают. Вот. Но здесь могут быть такие моменты, что кто-то из вас этого сделать не сможет. Ну, допустим, описание, работу сделает и описание, а разместить в РИНЦ и в ResearchGate не сможет по каким-либо причинам. Какие причины могут быть? Значит, в РИНЦ единственная причина, что вы очень долго там будете с этим заниматься этим вопросом. Ну то есть я вот сейчас вам говорю, что нужно прямо вот сейчас начать регистрироваться.
Ребят, вы спрашиваете меня голосом. Потому что вот то, что вы там в чате пишете, я вот сейчас посмотрел. Я же смотрю на экран, показываю вам пособие там, эту инструкцию.
Значит, для того, чтобы в РИНЦ зарегистрироваться, ничего не нужно. Просто заходите вот в этот РИНЦ и начинаете регистрироваться. Чтобы туда зайти, для этого нужно набрать, сейчас я вам покажу. Вот этот адрес. Вот я вам ссылочку вам кину. Вот эту, ну, elibrary.ru, то есть электронная библиотека и всё. Default - это вот не обязательно, вот это вот elibrary. И всё. И начинаете там регистрацию. Но единственное что, что при регистрации ещё нужно, там карточка будет регистрация, которую вы будете заполнять. В этой карточке, в нижней части слева есть место, где можно птичку поставить. Там будет вопрос такой: регистрироваться ли в системе Science Index? И птичку можно поставить. Там нужно ставить птичку. И регистрироваться в системе Science Index. Почему? Потому что если вы в ней не зарегистрируетесь в этой системе, то вы не получите тогда спин-код. А спин-код - это вот этот персональный код, Science Personal Index, ну то есть персональный идентификатор, короче говоря, в этом РИНЦе. РИНЦ - это Российский индекс научного цитирования. Вот этого вы не получите. Это присваивается только в том случае, если вы регистрируетесь не только в РИНЦ, но и в системе Science Index, вот она, система эта. А когда вы получите такой спин-код, это может, если вы сегодня вы начнёте регистрироваться в РИНЦе в этом, то, может быть, через дней 10-20 вы получите. Где-то через, ну, через одну, две, три недели, примерно так. То есть это не очень быстро получается, понимаете? Хотя заполняете вы быстро. Вы это всё быстро заполните, это нетрудно заполнить. А потом будете ждать, понимаете, пока вам будет присвоен этот этот индекс. Когда он будет присвоен, тогда вы выходите вот на эту ссылочку. Я сейчас рассказываю, как регистрироваться в системе РИНЦ, Science Index и заключить договор. Как найти, откуда я взял эту ссылочку, ребята? Вот откуда я вообще её знаю? Вот смотрите откуда. Вот я выхожу в elibrary, потом нажимаю автором, смотрите, автором. Потом вот так пролистываю и нахожу здесь размещение непериодических изданий, видите? Нажимаю, прокручиваю, и вот здесь вот есть система заключения договоров. Вот заходите в эту систему заключения договоров и начинаете заключать договор на физическое лицо. Не на организацию, а на физическое лицо договор заключить, на вас. После этого вы сможете размещать публикации, которые размещены в ResearchGate, а также свои какие-то монографии, учебные пособия, то есть непериодические публикации.
* **Процедура регистрации в ResearchGate и требование корпоративной почты**
Для того, чтобы разместить в ResearchGate, надо в ResearchGate зарегистрироваться. В ResearchGate зарегистрироваться несложно.
Мы, скорее всего, не сможем зарегистрироваться в ResearchGate, да?
А в ResearchGate, чтобы зарегистрироваться, там всё это очень просто делается, но нужно иметь корпоративный адрес электронной почты. То есть такой адрес почты, который заканчивается на домене КубГАУ, короче. Вот. Ну или edu.kubsa.ru, вот этот вот домен, или kubsa.ru, или edu.kubsa.ru. Edu.kubsa.ru - это, это гугловская почта. Но заведённая от имени университета. Вот. Так вот, ну, Google, Gmail, короче, это почта, на самом деле. Просто там домен другой указан.
Для того, чтобы в ResearchGate зарегистрироваться, вам надо обратиться к своему руководству студенческому там, ну, старосте, видимо, пойти, куратору. Куратор там, я не знаю, наверное, в деканат, может быть, обратиться, не знаю. В общем, нужно пытаться стучаться, выяснять, как вам можно получить корпоративный адрес электронной почты. Или, или, слушайте внимательно, ребята, или не получить его, а хотя бы, чтобы они подтвердили вам. Его выдаёт этот адрес, выдаёт Центр информационных технологий. Значит, есть университеты, скажем, вот КГУ, где я тоже преподаю, в КГУ. Вот. Вот смотрите. Вот КГУ. Вот студент Дмитрий Немировский. А вот адрес его корпоративной электронной почты. Вот это. Надеюсь, вам видно. Буковка S, потом номер, потом домен edu.kubsu.ru. А у нас такой же точно домен есть для преподавателей edu.kubsa.ru. Зарегистрировать там студентов вообще никаких проблем нет, понимаете? Вообще никаких проблем. Единственное - это работа просто. А у преподавателей такие же точно адреса, только там написано П впереди буковка. Вот. То есть вам надо вот завести адрес электронной почты, после этого вы в ResearchGate регистрируетесь там, ну, буквально там за 5 минут.
* **Важность регистрации для портфолио и навыков**
Значит, что это даёт вам? Это вам даёт две вещи очень важные. Первое - у вас пополняется портфолио. То есть это ваши достижения научные. И вы, я вам даю технологию. Знаете, вот как говорят, что есть возможность накормить человека, а есть возможность дать ему удочку. То есть можно дать человеку рыбу, если он голодный, а можно дать ему удочку, научить ловить рыбу. Вот. И так вот я сейчас вас именно не даю вам рыбу, а я вам даю удочку, понимаете? То есть я вам рассказываю, как вообще можно проводить научные исследования, как их оформлять бесплатно. Только от вас это зависит, чтобы вы этому научились. То есть вы сейчас вот, если будете делать то, что я вам говорю, то через некоторое время у вас появится бесплатная возможность публикации с присвоением DOI, размещённых в РИНЦ. Я вам могу сказать, что мне постоянно приходят в спам такое, нежелательная реклама. И среди неё, ну, всё-таки адрес размещён на сайте, в журнале у меня, то есть я такой довольно известный. И предлагается публиковать статьи в различных конференциях, там, сборниках конференций, от 65 рублей страница до 200. Значит, если с присвоением DOI, то обычно около 200 руб. страница стоит. То есть публикация РИНЦ с присвоением DOI стоит где-то около 200 руб. То есть получается так, что если вы хотите опубликовать статью, то там 1000, полторы, две придётся заплатить, понимаете, как минимум. А бывает и побольше. Это такая вот реклама самая-самая такая ширпотребная, дешёвая приходит. Я вам рассказываю, что это можно делать бесплатно, когда вам захочется. То есть вот вы сегодня захотели, взяли сегодня, провели исследование, описали его. А инструменты исследования тоже я вам даю, это система AIDOS. Вот. Вы показывал уже образцы, примеры этих исследований. Ещё вот сейчас мы будем смотреть, когда будем темы выбирать, всё я сейчас буду вам показывать, как это можно делать. Вот. И вы можете в любой момент провести исследование. Информация достаточна для этого, более чем достаточна в интернете. И описать его по шаблону и разместить в ResearchGate, у вас появляется публикация международная с DOI, Digital Object Identifier, то есть с международным идентификатором цифровым вашей публикации. И такие публикации молча РИНЦ размещает на своей платформе. Потому что это довольно престижная публикация международная. Вот. После этого у вас появляется публикация в списке ваших работ. И вы можете эту публикацию при защитах указывать в списке в автореферате, например. В автореферате там есть, когда диссертации защищаются, там есть несколько разделов. Первый раздел - это публикации в изданиях, зарегистрированных в Web of Science и Scopus. Потом монографии идут. Потом идут публикации, зарегистрированные в изданиях перечня ВАК, опубликованные в изданиях, входящих в перечень ВАК. А потом идут все остальные публикации. Так вот во всех остальных, там могут быть и публикации вот такие, как я сейчас вам сказал. Это публикации с присвоением DOI, входящие в РИНЦ. Вот там они могут быть. То есть таких публикаций, пока вы будете учиться, вы вполне можете там две, три, там, четыре, там сделать. А я видел студентов, у которых около тридцати публикаций уже. Ну, правда, это магистранты были. Вот. То есть уже это прямо вот можно защищаться, понимаете? Диссертацию можно уже выходить на защиту по числу публикаций.

**6. Оценка и соавторство**

* **Критерии оценки ("сам экзамен")**
Вот такие дела. Если вы всё это делаете, разрабатываете, размещаете в AIDOS-облако, в ResearchGate, в РИНЦ... В AIDOS-облаке со мной вместе или я сам размещу. Вот. То это тогда отличная оценка. Если вы в РИНЦ не можете разместить, тогда хорошо. Вот. Только ResearchGate и AIDOS-облако. Если только в AIDOS-облаке, это проще всего. Но я размещаю только то, что там как бы того стоит. Ну, как бы, более-менее нормальное качество работы. Тогда удовлетворительно. А если, значит, ничего этого нет, ну тогда, э-э, ну как, таких не должно быть, не должно быть таких студентов вообще. Но если будут, ну тогда будем смотреть, что там у них по аттестации, там, в общем, короче, это похоже, что как двойка, по идее. Вот. Ну двоек сейчас, наверное, ставить мы не будем, и вообще я их не люблю ставить, но, в общем, короче говоря, нехорошо получается, если вообще ничего нет. Вот, должно быть. Вы учились, осваивали, должно всё быть у вас.
А можно не шутя на сам экзамен?
Значит, сам экзамен - это вот это и есть, когда вы всё это вот вы сделаете, что я вам перечислил, вот эти все моменты. Тут вот и есть сам экзамен. То есть это означает, что я ничего спрашивать не буду, потому что вы полностью подтвердили свои компетенции, знания, умения, навыки, всё сделали. Значит, значит, вы во всём разбираетесь, я ставлю тогда сам экзамен.
Вот этот человек сделал эту работу, получил сам экзамен. Понятно, что такое сам экзамен? Даниил? Даниил? Скажи: "Понятно". Или скажи: "Непонятно вообще, что такое сам экзамен?" Что ты молчишь? Давай, говори. Понятно или нет? Да, понятно.
* **Вопрос соавторства с преподавателем**
Вот здесь я хочу, ребята, ваше внимание на что акцентировать? Что если вы не зарегистрируетесь в ResearchGate, то тогда мне придётся там размещать ваши работы, если вы их сделаете. Если вы не зарегистрируетесь в РИНЦ, то мне придётся размещать ваши работы, если вы их сделаете. А для того, чтобы размещать работы ваши в ResearchGate, в РИНЦ, для этого нужно быть соавтором. Поэтому получится так, что, ну, у меня-то есть основания быть соавтором, потому что я разработал технологию, с которой вы будете работать, и все учебно-методические материалы, и программное обеспечение, всё, что тут вообще есть, всё это я разработал. Вот. И предоставил вам. Вот. Но, то есть у меня полное основание быть соавтором, но я бы не хотел, если честно, уже быть соавтором. Почему? Потому что у меня слишком много работ уже. Ну, может быть, конечно, их и не бывает слишком много. Ну, в общем, здесь вот у меня написано 625 работ. Сейчас я посчитал, у меня где-то больше 630 работ сейчас. Я просто не отражаю их в списке, потому что сейчас ещё одна выйдет работа, тогда я обновлю список. Будет где-то 632 вот здесь вот работы, понимаете? 38 монографий, 27 учебных пособий, 30 патентов. Ну то есть 280 с лишним публикаций в изданиях, входящих в перечень ВАК. Ну то есть мне, в принципе, достаточно этого, понимаете? То есть нет никакой необходимости ещё иметь публикации. Вот. Но если вы сами не сможете разместить свои публикации в РИНЦ, в ResearchGate, то это придётся мне делать. Ну тогда, если они того стоят, я там не против поставить свою фамилию. Вот. И разместить их.

**7. Возможность коллаборации**

* **Работа в парах**
Вопрос часто возникает у студентов, а можно не одному, а вдвоём работу выполнить? С кем-то? То есть вот не один автор-студент, а два студента-автора работы? Вот смотрите, работа Луценко, Аладинский, Смирнова. Или вот Захаров, Нехай, Захаров, Поспелов. Видите? Или вот Мотылец, Ульянов. Ну, а впереди я стою. Это что значит? Да можно, ребята, можно вдвоём выполнить работу, вдвоём можно. То есть я вас, может быть, обрадую этим. Ну, в общем, короче, вы можете выполнить работу вдвоём, одну. И у вас обоих будет сам экзамен. Но втроём - это уже перебор. Или всей группой одну работу, если вы выполните, это уже будет слишком.

**8. Демонстрация работы с данными (CSV)**

* **Импорт CSV в AIDOS**
CSV файлы, ребята, я сейчас в новой версии системы, которая вот буквально на днях размещена, вот набор данных CSV. Значит, в Экселе считывается он. Значит, реализовал режим, обеспечивающий непосредственный ввод данных из CSV файлов. То есть раньше в системе AIDOS были режимы, которые обеспечивали ввод из экселевских файлов, старого и нового Экселя, и dBase файлов, DBF. А теперь ещё есть и CSV. Там раньше это было заявлено. Это что такое CSV файл? Это текстовый файл, ну, форматированный. В нём точно так же, как в Экселевском, есть поля, колонки. И в этих колонках могут быть числовые и текстовые данные. Вот. Такой файл просто вводится в систему непосредственно. Причём очень большого размера. Этот конвертер, который я сделал на днях, он работает, я его апробировал, проверил на файле, в котором было 4 млн записей строк.
Чтобы сделать модель из CSV файла, что нужно сделать, ребят? Вы должны войти в папочку системы AIDOS, открыть папочку AID Data, папочку InData открыть, удалить там всё, туда записать CSV файл и дать ему имя InData, такое же, как имя папочки. Иначе система скажет, что там такого файла нет. Вот у нас есть файл InData, видите? В папочке, значит, AID Data. Вот путь, вот путь на систему у вас может быть другой, но потом там вот AID Data, InData и файл этот.
* **Запуск анализа в AIDOS**
Теперь берём, что мы делаем? Я не очень уверен, что будет сейчас работать нормально, потому что запущен очень тяжёлый для компьютера процесс, он перегруженный сейчас. Ну попробуем. Вроде работает. Вот. И входим в систему. Вот это окошко системы. Стираем все приложения, какие там были. Заходим в режим 2 3 2 2, где мы вводили экселевскую таблицу. А здесь, смотрите, ребята, указываем, что мы используем CSV файл. Видите, да? И значит, здесь вот описано, какой у него стандарт этого CSV файла. Да, вот ещё что. Значит, мы можем выйти в Excel, ну, желательно такой, более-менее новый Excel. Вот. Ну у меня 2010, это так достаточно новый Excel. Вот. И мы выбираем вкладочку Данные, смотрите, Данные. Вот из текста. И указываем путь на CSV файл. Для чего это нужно? Для того, чтобы посмотреть, какие у него будут поля у этого файла. И смотрим, какой разделитель между полями. Вот у него разделитель запятая, сразу раз колоночки получились. Вот этот файл сейчас ведётся. Вот вёлся он. Он в Excel, кстати, конвертер Экселя вводит его с ошибками этот файл. Система AIDOS без ошибок вводит. И вот у нас, видите, есть колоночка рейтинг. Вот эта колоночка, она у нас является какой по номеру? Сейчас мы посчитаем. Раз, два, пятнадцатый. Вот мы и записываем. В колоночке со второй по четырнадцатую - это у нас описательные шкалы, а пятнадцатая колоночка - это классификационная шкала. То есть мы хотим узнать, как вот это всё влияет на рейтинг в той области, в которой этот эти данные собраны. Вот мы, значит, сейчас и будем это делать. Задаём здесь вот, что мы используем CSV файл для конвертации его в dBase, указываем, что у нас классификационная шкала пятнадцатая, а описательные со второй по четырнадцатую. И О'кей нажимаем. Происходит преобразование этого файла в dBase. Вот сейчас это видно, как это происходит, видите? Значит, сначала он просматривается и определяются имена полей, длина поля и тип данных поля. А потом создаётся файл соответствующей структурой и числовыми и текстовыми полями, и вводится туда данные. 28 секунд это удовольствие заняло у нас. Вот. И начинается процесс уже обработки dBase файла в системе AIDOS. Сейчас это всё происходит немножко так с тормозами, потому что очень трудоёмкий процесс идёт в вычислительном плане. И определено у нас, что у нас 1314 классов, 4363 значения фактора или признака. Вот. Ну, значит, числовых шкал у нас классификационных вообще нету, описательных три у нас числовые шкалы с небольшим числом градаций. Выйти на создание модели. Создаются справочники классификационных и описательных шкал и градаций. Кодируем исходные данные с помощью этих справочников, получаем обучающую выборку. 16.000 примеров. Каждый из этих примеров вот у нас описан какими-то кодами признаков. И уже созданы модели.
* **Проверка достоверности модели**
Теперь мы проверяем эти модели на достоверность. Это делается отдельно. Только верификация модели. Почему? 16.000 объектов - это может около часа идентифицироваться. Поэтому я беру здесь вот делаю, допустим, не 16.000, а 300, например, там объектов беру. Только верификацию провожу. И указываю, чтобы выбрать наиболее достоверные результаты. То есть я провожу верификацию на трёхстах объектах обучающей выборки, а не на 16.000. Вот идёт распознавание. То есть все 300 объектов, каждый из них сравнивается с 1300 классов, ребят, по 1300 признакам. То есть эта задача огромной вычислительной мощности требует. Каждый из объектов сравнивается с каждым из 1300 классов, а для каждого сравнения нужно 1300 сложений произведений сделать. Вот. Ну на графическом процессоре это происходит очень быстро. Потом создаётся 10 выходных форм. И самое главное, что создаётся форма по, описывающая достоверность модели по F-критерию Вальда-Бергена и его обобщению. Вы помните про это, что я вам про это рассказывал, ребята? Или уже забыли начисто? А? Ну это одна модель посчиталась, а их там 10. Поэтому это всё равно займёт время.

**9. Заключение и дальнейшие шаги**

* **Подведение итогов и инструкции**
Вы пока можете вопросы позадавать. То есть я сейчас вам показываю, как реализуется пункт пятый. Только вы должны сами это делать. Находить данные. Я советую в Кагле всё-таки. Вот. Данные не очень большой размерности, с числовыми шкалами. Вот. Чтобы было два CSV файла. В основном с числовой информацией. И ставите задачу, решаете её, смотрите достоверность. Ну если получилось неплохо, шлёте мне это и просите, чтобы я утвердил эту тему. Я её утверждаю. После этого вы можете делать спокойно вот это описание. И если можете, если у вас уже есть регистрация в ResearchGate и РИНЦ, там размещать. Если при этом будут возникать вопросы, будете меня спрашивать, я буду подсказывать. Вот то, что я сейчас вам рассказываю, примерно это делают ребята из КГУ, вот этот Усов Павел и Немировский Дмитрий. Примерно вот это и делают. Вот. И у них получается. Елисей Негодин он в ResearchGate уже разместил свои материалы. Староста. То есть есть ребята, которые прямо вот это делают то, что я им говорю. Другие тоже делают, но с тормозами, замедленно, в замедленном темпе.
Так, ребята, у нас на этом занятие заканчивается. То есть вы сейчас начинаете непосредственно выполнять то, что я вам сказал: регистрироваться в ResearchGate и РИНЦ. Вот, и искать тему. И пробовать, как она работает, получается, не получается, насколько это получается. И мне об этом сообщать. Ну, это вот у нас затягивается процесс, поэтому я его прерву просто и всё.
* **Завершение**
И на этом, ребята, конец занятия. Всего самого хорошего. До свидания.
До свидания.
Вопросы если есть, задавайте. А то я сейчас прерву. Вопросы есть какие-нибудь, нет у вас? В общем, у нас дальнейшее занятие, ребят, посвящено выполнению этого задания. Это задание комплексное, оно полностью проверяет все ваши знания, умения, навыки по дисциплине. И по его результатам выставляется оценка. Всё, самого хорошего. До свидания.